

BLACK BOXES

*Julie A. Seaman**

ABSTRACT

The metaphor of the black box has often been used to describe the qualities of the human mind; likewise, the jury box is frequently referred to as a black box. In both contexts, the metaphor is apt because of the inscrutability of the process that gives rise to the outputs that emanate from each. Recent advances in brain imaging techniques have now begun to crack open the black box that is the human mind by illuminating the physical manifestations—the “neural correlates”—of a wide range of cognitive processes. In particular, research into the neural correlates of deception presents the genuine prospect of a reliable, forensically practicable lie detector within the foreseeable future. This Article proceeds in the nature of a thought experiment to explore the ramifications for the jury system of a highly reliable lie detection technique. In particular, I suggest that opening the black box of the mind would have the effect of opening the black box of the jury room.

Conventional wisdom has it that the jury’s primary—if not singular—function is to determine the historical facts of the case. Yet it is clear that in addition to finding facts, juries also operate in the much more controversial realm of making law. At its extreme, this lawmaking role may result in jury nullification, whereby the jury issues a verdict intentionally contrary to the law as instructed by the court, applied to the facts as found by the jury. Whereas the jury’s power to nullify is well-settled, its right to nullify is highly contested. Thus, much of the scholarly and judicial discussion has focused on the issue of whether the jury may or must be instructed that it has the ability to return a verdict contrary to the applicable law. Though scholars are divided, courts have uniformly held that juries should not be told of their power to nullify.

* Associate Professor, Emory University School of Law. J.D., Harvard Law School. For helpful comments on earlier drafts, I thank Dorothy Brown, Michael Kang, Kay Levine, Jonathan Nash, Victoria Nourse, Robert Schapiro, and Peter Weitzner; in addition, Erin Murphy and Jennifer Mnookin shared their invaluable questions and insights in informal conversation. I am grateful to have had the opportunity to present these ideas at the 2008 Randolph W. Throver Symposium at the Emory University School of Law, as well as at the Neuroscience, Law and Government Symposium held at the University of Akron School of Law.

To the extent that brain-imaging lie detection techniques (along with other technological advances in forensics) diminish the need for jury fact-finding, the jury's lawmaking role would become more transparent to the public and, perhaps equally important, to the jury itself. In cases in which the facts were clear, the possibility and the actuality of nullification also would become clear. Thus would arise the questions: Is the black box quality of jury decision making integral to the nature of the jury system itself? Would opening the black box destroy it? Should even highly accurate lie detection evidence be excluded to preserve the black box nature of jury decision making? This Article offers a framework within which to begin to think about these questions.

INTRODUCTION	429
I. INSIDE THE FIRST BLACK BOX: WATCHING THE BRAIN THINK	441
A. <i>The Brain on Lies—an Overview</i>	443
B. <i>The Brain-Imaging Studies</i>	445
1. <i>“Guilty Knowledge” Playing Card Studies</i>	447
2. <i>Increasing Ecological Validity</i>	449
3. <i>Malingering</i>	450
4. <i>Autobiographical Lies</i>	451
C. <i>Recent Advances</i>	452
II. INSIDE THE SECOND BLACK BOX: WATCHING THE JURY DECIDE	456
A. <i>Shutting the Lid: Protecting the Jury as Sole Fact Finder</i>	457
B. <i>Pressure Building Outside the Box</i>	466
C. <i>The Box Splits Open: What Remains of the Jury’s Role?</i>	475
III. LEGISLATING INSIDE THE BLACK BOX: JURY NULLIFICATION	478
CONCLUSION: PANDORA’S JURY BOX	487

*When a machine runs efficiently, when a matter of fact is settled, one need focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become.*¹

*The jury is the justice system's black box; historical information enters and legal conclusions are produced. The methods and reasons for arriving at those conclusions are unknown in ordinary review. We decline to hypothesize about the jury's reasoning.*²

INTRODUCTION

This Article considers black boxes and their place in the resolution of legal disputes. The phenomenon of jury decision making in cases involving complex science is awash in black box metaphors. The use of science in the courtroom—and especially very complex science well beyond the understanding of typical legal actors³—presents something like a nesting of black boxes, one inside the other. First, there is what many have called the black box of the human mind.⁴ Next, there is that classic black box of the litigation process: the jury.⁵ And at the most general level, there is the paradoxical black box of modern science and technology, as described in the

¹ BRUNO LATOUR, PANDORA'S HOPE: ESSAYS ON THE REALITY OF SCIENCE STUDIES 304 (1999) (defining the term "blackboxing" as it is used in the sociology of science).

² *Children Int'l v. Ammon Painting Co.*, 215 S.W.3d 194, 200 (Mo. Ct. App. 2006).

³ See Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CAL. L. REV. 721, 729 (2007) (distinguishing "second generation" forensic techniques as "intuitively inaccessible to laypersons, even while the results of such methods are typically viewed as highly reliable," and as typically using "technically sophisticated methods" that are "costly, elaborate, and incapable of ready reproduction").

⁴ "[U]ntil recently, science could not tell us much about how the human brain works. The brain was a 'black box,' a mass of messy tissue into which the reports of our senses disappeared and out of which behavioral decisions emerged." Oliver R. Goodenough, *Mapping Cortical Areas Associated with Legal Reasoning and Moral Intuition*, 41 JURIMETRICS 429, 431 (2001); see SUSAN ALLPORT, EXPLORERS OF THE BLACK BOX: THE SEARCH FOR THE CELLULAR BASIS OF MEMORY (1986); D.W. HAMLYN, IN AND OUT OF THE BLACK BOX: ON THE PHILOSOPHY OF COGNITION (1990); Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511 (2004).

⁵ See, e.g., Vivian Berger, *"Black Box Decisions" on Life or Death—If They're Arbitrary, Don't Blame the Jury: A Reply to Judge Patrick Higginbotham*, 41 CASE W. RES. L. REV. 1067 (1991); Daniel R. Cahoy & Min Ding, *Using Experimental Economics to Peek into the "Black Box" of Jury Behavior: A Proposal for Jury Research Reform*, 14 S. CAL. INTERDISC. L.J. 31 (2004); Valerie P. Hans, *Inside the Black Box: Comment on Diamond and Vidmar*, 87 VA. L. REV. 1917 (2001); Kimberly A. Moore, *Judges, Juries, and Patent Cases—An Empirical Peek Inside the Black Box*, 99 MICH. L. REV. 365 (2000).

first epigraph above. Every so often, the opportunity arises to peer inside a black box; the choice whether to do so seems to depend on a complex balancing of curiosity and fear. What if it should turn out that, as with Pandora's box, a host of evils is let loose that, once released, cannot be undone?⁶

Rapid advances in neuroscience over the past few years have raised the genuine prospect that the inner workings of the human brain might soon become visible. Already, scientists using powerful brain imaging technologies have begun to distinguish lying brains from truthful brains,⁷ false memories from true memories,⁸ and “malingerers” from those experiencing genuine symptoms.⁹ Research in cognitive neuroscience¹⁰ has uncovered distinctive regions of the brain that are differentially activated when, for example, subjects view photographs of out-group versus in-group members,¹¹ when they apply word-based rules versus moral intuitions of justice,¹² and when they experience feelings of disgust¹³ or romantic love.¹⁴ The processes of decision making, including those outside conscious awareness,¹⁵ have become increasingly

⁶ The Second Circuit has likened the prospect of the admissibility of polygraph evidence to a “legal Pandora’s box.” *United States v. Kwong*, 69 F.3d 663, 668–69 (2d Cir. 1995).

⁷ See *infra* notes 62–123 and accompanying text.

⁸ See Scott D. Slotnick & Daniel L. Schacter, *A Sensory Signature that Distinguishes True from False Memories*, 7 NATURE NEUROSCI. 664 (2004).

⁹ See sources cited *infra* note 85.

¹⁰ “Cognitive neuroscience is an investigational field that seeks to understand how human sensory systems, motor systems, attention, memory, language, higher cognitive functions, emotions, and even consciousness arise from the structure and function of the brain. . . . [It] has been described as a ‘bridging discipline’—between biology and neuroscience, on the one hand, and cognitive science and psychology, on the other.” O. Carter Snead, *Neuroimaging and the “Complexity” of Capital Punishment*, 82 N.Y.U. L. REV. 1265, 1273 (2007) (citing HOWARD S. KIRSHNER, *BEHAVIORAL NEUROLOGY: PRACTICAL SCIENCE OF MIND AND BRAIN* 5 (2d ed. 2002); *THE COGNITIVE NEUROSCIENCES III* (Michael S. Gazzaniga ed., 2004)).

¹¹ See W.A. Cunningham et al., *Separable Neural Components in the Processing of Black and White Faces*, 15 PSYCHOL. SCI. 806 (2004); Allen J. Hart et al., *Differential Response in the Human Amygdala to Racial Outgroup Versus Ingroup Face Stimuli*, 11 NEUROREPORT 2351 (2000); Elizabeth A. Phelps et al., *Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation*, 12 J. COGNITIVE NEUROSCI. 729 (2000).

¹² See Goodenough, *supra* note 4, at 442 n.64; Johannes Schultz et al., *Cortical Regions Associated with the Sense of Justice and with Legal Rules*, 13 NEUROIMAGE S473 (2001).

¹³ See Andrew J. Calder et al., *Neuropsychology of Fear and Loathing*, 2 NATURE REV. NEUROSCI. 352 (2001).

¹⁴ See A. Bartels & S. Zeki, *The Neural Basis of Romantic Love*, 11 NEUROREPORT 3829 (2000).

¹⁵ See generally SOCIAL PSYCHOLOGY AND THE UNCONSCIOUS: THE AUTOMATICITY OF HIGHER MENTAL PROCESSES 51 (John A. Bargh ed., 2007). See also John A. Bargh & Tanya L. Chartrand, *The Unbearable Automaticity of Being*, 54 AM. PSYCHOLOGIST 462, 465 (1999) (collecting and describing evidence of unconscious processes across a range of mental activities, and stating that “most of our day-to-day actions, motivations, judgments, and emotions are not the products of conscious choice and guidance, but must be

visible. Perhaps most amazing—and most disconcerting—is that scientists in several recent experiments have been able to determine with a high degree of accuracy what subjects were thinking about or looking at simply by analyzing scans of their brains.¹⁶ Though certainly years, and perhaps decades, remain before such research reaches a level of reliability sufficient for courtroom use,¹⁷ it seems fair to say that the future approaches the doorstep—if it is not already poised at the threshold.¹⁸

driven instead by mental processes put into operation directly by environmental features and events”); Jacob R. Waldbauer & Michael S. Gazzaniga, *The Divergence of Neuroscience and Law*, 41 JURIMETRICS 357, 362 (2001) (“The neuroscientific view is that human behavior is the result of rule-following by our automatic brains.”).

¹⁶ See Kendrick N. Kay et al., *Identifying Natural Images from Human Brain Activity*, 452 NATURE 352 (2008); Tom M. Mitchell et al., *Predicting Human Brain Activity Associated with the Meanings of Nouns*, 320 SCIENCE 1191 (2008); Svetlana V. Shinkareva et al., *Using fMRI Brain Activation to Identify Cognitive States Associated with Perception of Tools and Dwellings*, 3 PLOS ONE e1394 (2008).

¹⁷ See Jane Campbell Moriarty, *Flickering Admissibility: Neuroimaging Evidence in the U.S. Courts*, 26 BEHAV. SCI. & L. 29 (2008) (canvassing cases in which neuroimaging evidence was considered, and showing that admissibility has thus far been limited to demonstrating brain trauma or disease). As Professor Moriarty notes, “There are no cases to date admitting fMRI evidence as proof of deception or truth-telling. Indeed, scholarly commentary explaining the shortcomings of the various studies would suggest that the evidence could not yet meet either the *Frye* general acceptance standard or the *Daubert* reliability standard.” *Id.* at 46. *But see* Sean A. Spence et al., ‘*Munchausen’s Syndrome by Proxy*’ or a ‘*Miscarriage of Justice*’? *An Initial Application of Functional Neuroimaging to the Question of Guilt Versus Innocence*, 23 EUR. PSYCHIATRY 309, 312 (2008) [hereinafter Spence et al., *Munchausen’s Syndrome*] (“To our knowledge, this is the first case described where fMRI or any other form of functional neuroimaging has been used to study truths and lies derived from a genuine ‘real life’ scenario, where the events described pertain to a serious forensic case.”). Recently, the *New York Times* reported on a murder case in India in which “a judge explicitly cited a scan as proof that the suspect’s brain held ‘experiential knowledge’ about the crime that only the killer could possess, sentencing her to life in prison.” Anand Giridharadas, *India’s Novel Use of Brain Scans in Courts Is Debated*, N.Y. TIMES, Sept. 15, 2008, at A10.

¹⁸ It is of course important to be alert to the human propensity to be overly optimistic—sometimes wildly so—about the potential for new scientific techniques to solve intractable problems. An examination of contemporaneous statements made at the dawn of the polygraph era is sobering: the machine was said to be “almost infallible in its detection of whether or not a person is guilty of a crime.” KERRY SEGRAVE, *LIE DETECTORS: A SOCIAL HISTORY* 10 (2004). An earlier precursor was said to “furnish[] a certain and objective criterion between truth and falsehood.” *Id.* at 15. In 1921, the *Boston Sunday Advertiser* stated that “[n]o matter how accomplished at ordinary deception a man may be, he cannot hope to deceive [polygraph inventor] Marston’s apparatus any more than a woman can humbug a weighing machine by lacing tightly and dressing in black.” *Id.* at 16. Some have cautioned that neuroimaging results are especially likely to evoke disproportionate expectations because of the tantalizing images used to model the results. See Paul Bloom, *Seduced by the Flickering Lights of the Brain*, SEED, June–July 2006, at 20, 22, available at http://www.seedmagazine.com/news/2006/06/seduced_by_the_flickering_ligh.php (arguing that fMRI images have captivated the media and the public beyond their actual worth); Laura Stephens Khoshbin & Shahram Khoshbin, *Imaging the Mind, Minding the Image: An Historical Introduction to Brain Imaging and the Law*, 33 AM. J.L. & MED. 171, 182 (2007) (stating that when courtroom evidence includes images of brain scans, the search for truth is apt to be distorted because “brain-images can be profoundly fascinating to view . . . [and] the visual allure of the images in combination with the endorsement of a neuroscientist may result in juries being too easily persuaded of their evidentiary value in a case”); Deena Skolnick Weisberg et al., *The Seductive*

With the ability to unlock the black box of the mind—for example, to know whether a witness’s testimony is truthful, deceptive, or simply mistaken—would come the prospect of a clearer view into that other black box: the jury box. Typically, a jury verdict in a criminal case¹⁹ is inscrutable; the jury performs its paradigmatic function as fact finder shrouded in secrecy, and it is impossible to say why or how the jury convicted or acquitted in any given case.²⁰ Indeed, the *Federal Rules of Evidence*, along with other doctrinal and procedural rules, work to ensure that this will be so.²¹ Observers can, and

Allure of Neuroscience Explanations, 20 J. COGNITIVE NEUROSCI. 470, 477 (2008); Jennifer Kulynych, Note, *Psychiatric Neuroimaging Evidence: A High-Tech Crystal Ball?*, 49 STAN. L. REV. 1249, 1251 (1997). I am mindful of this danger and recognize that it is fully possible that a reliable lie detector will never be developed. However, I believe that the current project is nonetheless useful for two reasons: First, given the very rapid advances in the field, it is at least conceivable that the technology will reach a stage of high accuracy. It is preferable to consider the potential implications of the technology in advance of that point. Second, purely as a thought experiment it is useful to ask what, if anything, would remain of the jury’s role in the criminal justice system were its fact-finding function to become obsolete. Such an inquiry reveals important insights about the institutional role of the jury and also about the potential importance of concealing its workings within a black box.

¹⁹ Though courts certainly exert greater control over the decision making of civil juries, jury decisions retain their black box quality to a great extent even in civil cases. Purely in terms of inscrutability, the single most important distinction is the more prevalent use of special verdicts in civil cases. See John D. Jackson, *Making Juries Accountable*, 50 AM. J. COMP. L. 477, 519 n.161 (2002). Compare FED. R. CIV. P. 49(a) (allowing special verdicts in federal civil cases at the discretion of the court), with *United States v. Spock*, 416 F.2d 165, 181 (1st Cir. 1969) (stating that jurors should be free “from judicial pressure, both contemporaneous and subsequent”).

²⁰ See, e.g., NEIL VIDMAR & VALERIE P. HANS, *AMERICAN JURIES: THE VERDICT* 228 (2007) [hereinafter VIDMAR & HANS, *AMERICAN JURIES*] (“Unless jurors speak forthrightly during jury selection, however, or admit they are disregarding the law during jury deliberation, it’s difficult to determine whether jury nullification is occurring and nearly impossible to police nullifying jurors.”).

²¹ “Undisputed procedures—the general verdict, the principle of noncoercion of jurors, and the inability to direct verdicts of conviction—ensured both nineteenth- and twentieth-century American juries the practical power to ‘acquit against instructions.’” Albert W. Alschuler & Andrew G. Deiss, *A Brief History of the Criminal Jury in the United States*, 61 U. CHI. L. REV. 867, 914 (1994) (citation omitted). Federal Rule of Evidence 606(b) prohibits jurors from testifying

as to any matter or statement occurring during the course of the jury’s deliberations or to the effect of anything upon that or any other juror’s mind or emotions as influencing the juror to assent to or dissent from the verdict or indictment or concerning the juror’s mental processes in connection therewith.

FED. R. EVID. 606(b). Hearsay statements and affidavits of jurors on such matters are also prohibited. *Id.* Though the rule does allow jurors to give evidence of “outside influence” or mistake, the Supreme Court has interpreted this category narrowly. See *Tanner v. United States*, 483 U.S. 107, 125 (1987) (holding that drug and alcohol use by jurors during trial was not “external influence” and therefore post-verdict testimony by jurors was barred by Federal Rule of Evidence 606(b)).

often do, speculate about the prevalence of jury nullification,²² but they can never say definitively whether a particular jury has acquitted despite the facts, or because of them.²³ Much of this ambiguity rests on the jury's unique and protected role in evaluating witness credibility; there always remains the possibility that even seemingly overwhelming evidence of guilt has been rejected by the jury as not credible.²⁴ However, if a machine finally were able to reveal whether a witness was telling the truth or lying,²⁵ it would become much more apparent—to the public and, perhaps more significantly, to the

²² A nullificatory verdict is most often defined as a jury's intentional acquittal contrary to the law as instructed by the court, applied to the facts as found by the jury. See *infra* note 227. Though there exists a large and growing body of empirical work on jury nullification, see, e.g., Irwin A. Horowitz et al., *Jury Nullification: Legal and Psychological Perspectives*, 66 BROOK. L. REV. 1207 (2001), it remains very difficult to determine whether a jury has nullified in any particular case. See *supra* note 20.

²³ For example, many jury nullification scholars, as well as media commentators, have remarked on the so-called "Bronx juries" phenomenon, whereby inner-city juries apparently acquit at higher rates than the national average. See Elissa Krauss & Martha Schulman, *The Myth of Black Juror Nullification: Racism Dressed Up in Jurisprudential Clothing*, 7 CORNELL J.L. & PUB. POL'Y 57, 61–62 (1997); cf. VIDMAR & HANS, AMERICAN JURIES, *supra* note 20, at 286–87 (discussing the "Bronx effect" in civil cases). While some view the higher acquittal rates as evidence of nullification, see Nancy S. Marder, *The Interplay of Race and False Claims of Jury Nullification*, 32 U. MICH. J.L. REFORM 285, 287–301 (1999) (discussing widespread reaction of press and public to two notorious jury acquittals), others argue that it is entirely plausible to view these outcomes as evidence-based, for example, because such jurors' life experiences cause them to be more suspicious of police testimony, e.g., JEFFREY ABRAMSON, WE, THE JURY xvii (2000); Kraus & Schulman, *supra*. Still others dispute that these juries in fact exhibit significantly higher rates of acquittal in comparable cases. See Theodore Eisenberg & Martin R. Wells, *Trial Outcomes and Demographics: Is There a Bronx Effect?*, 80 TEX. L. REV. 1839, 1869–70 (2002). For further discussion of this point, see *infra* notes 227–35 and accompanying text.

²⁴ See, e.g., *United States v. Johnson*, 114 F.3d 808 (8th Cir. 1997). The court noted that, although the defendant had presented ample evidence from which the jury could have found him not guilty, "the law is well established that it is the jury's function to evaluate the credibility of witnesses." *Id.* at 812. The court further stated:

Given the jury's opportunity to view the witnesses and evaluate the witnesses' possible motives or biases, the jury was uniquely situated to determine whose testimony was credible," it would "not usurp the function of the jury to determine credibility, absent a finding that no reasonable person could have reached a conclusion the defendant was guilty beyond a reasonable doubt.

Id. at 813.

²⁵ "The search for truth, and its separation from lies, has been one of humankind's preoccupations since time began." SEGRAVE, *supra* note 18, at 183; see *id.* at 3–9 (describing ancient techniques of lie detection from various parts of the world, including submersion in water, boiling water, fire and hot irons, food ordeals, the "Chinese donkey tail test," the "fulcrum test," and an ancient Greek method of testing the pulse). In 1923, Dean Wigmore surmised that "[i]f there is ever devised a psychological test for the valuation of witnesses, the law will run to meet it." 2 JOHN HENRY WIGMORE, A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW § 875 (2d ed. 1923). Interestingly, the polygraph as originally developed was referred to as a "black box," presumably based on its physical appearance. See SEGRAVE, *supra* note 18, at 10 (exploring the development of the polygraph in a chapter entitled "The Black Box Takes Shape, 1870 to 1929").

jurors themselves²⁶—that the function of the jury is not limited to simple fact-finding.²⁷

Ultimately, then, the prospect of a highly reliable lie detector raises fundamental questions about the role of the jury in our civil and criminal justice systems, and indeed about the purpose of the jury trial itself.²⁸ If the fact-finding function of the jury were to (continue to) diminish over time due to scientific advances—including, for example, the development of a noninvasive, extremely reliable, and highly accurate lie detector—would trial by jury ultimately become an anachronism like the medieval trials by ordeal and by battle?²⁹ The answer to this question, as well as to questions about how

²⁶ The debate over jury nullification often centers on the question of jurors' knowledge of their ability to nullify. "The biggest academic debate over jury nullification, and one that has spilled over into the public realm, is whether juries should be instructed about their power to nullify in order to serve the interests of justice." VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 236. The overwhelming majority of courts that have considered the question have held that defendants are not entitled to a jury instruction informing jurors of their power to nullify the law. *See, e.g.*, *United States v. Wiley*, 503 F.2d 106, 107 (8th Cir. 1974); *United States v. Dougherty*, 473 F.2d 1113, 1130–37 (D.C. Cir. 1972); *United States v. Dellinger*, 472 F.2d 340, 408 (7th Cir. 1972); *United States v. Simpson*, 460 F.2d 515, 519–20 (9th Cir. 1972); *United States v. Boardman*, 419 F.2d 110, 116 (1st Cir. 1969); *United States v. Moylan*, 417 F.2d 1002, 1005–09 (4th Cir. 1969); *see also* discussion *infra* notes 237–49 and accompanying text. Indeed, several courts have excused jurors or declared mistrials when it became known that jurors were aware of their power to nullify the law. *See* Nancy S. Marder, *The Myth of the Nullifying Jury*, 93 Nw. U. L. REV. 877, 952 (1999) (gathering cases illustrating this point).

²⁷ This process may already be underway in cases involving DNA evidence because of its extremely high reliability and probative value in many cases. I thank Professors Jennifer Mnookin and Erin Murphy for raising this point in conversation. *See also* Michael S. Pardo, *Neuroscience Evidence, Legal Culture, and Criminal Procedure*, 33 AM. J. CRIM. L. 301, 305 (2006) (arguing that scientific evidence regarding witness credibility is in principle no different from other types of scientific evidence, such as DNA identification evidence, in terms of its effect on the role of the jury); Ric Simmons, *Conquering the Province of the Jury: Expert Testimony and the Professionalization of Fact-Finding*, 74 U. CIN. L. REV. 1013, 1053–57 (2006) (rejecting the argument that "questions of credibility are qualitatively different from the other facts put before a jury").

²⁸ *See generally* Symposium, *Truth & Its Rivals: The Goals of Evidence Law*, 49 HASTINGS L.J. 289–357 (1998). *See also* Chris William Sanchirico, *Character Evidence and the Object of Trial*, 101 COLUM. L. REV. 1227, 1230–31 (2001) (challenging the standard search-for-truth explanation with respect to the rules governing admissibility of character evidence, and arguing that these rules can better be understood as incentive-setting devices).

²⁹ *See* VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 21–27 (discussing the historical origins of the modern jury, including medieval methods of proof of trial by battle, ordeal, compurgation, and witnesses); Edward L. Rubin, *Trial by Battle. Trial by Argument.*, 56 ARK. L. REV. 261 (2003). As one scholar recently pondered, "What role would remain for the jury if scientists agree that credibility determinations could be made much more accurately by machines than by people? What would trials look like? Indeed, would there be any need for trials at all?" Mark Pettit, Jr., *FMRI and BF Meet FRE: Brain Imaging and the Federal Rules of Evidence*, 33 AM. J.L. & MED. 319, 333–34 (2007); *see also* *Cassamassina v. State*, 657 So.2d 906, 913 (Fla. Dist. Ct. App. 1995) (en banc) (Dauksch, J., dissenting) ("Of course, when such an instrument or method [to objectively determine the truth] is found then it won't be used to aid the jury, it can supplant it. Why have

the system should respond to such scientific advances, depends in large part on how the jury's role and purpose are conceived. If the jury is viewed primarily as a fact finder and lie detector,³⁰ then technologies that provide more accurate fact-finding and better lie detection arguably should replace a jury system that, increasingly, appears disturbingly inaccurate and prone to error.³¹ If, however, the jury serves unique and critical functions in addition to determining the historical facts at issue in the particular case, then the changes in technology might instead have the effect of bringing these more hidden functions into

a verdict . . . when a machine is better?"); SEGRAVE, *supra* note 18, at 30 (describing a commentary in the 1930s on the use of the polygraph, which stated: "If the lie detector is an absolutely infallible instrument, it seems to me that criminal courts are no longer necessary" (quoting *The Liars*, in 13 NEW STATESMAN AND NATION 362-63 (1937))).

³⁰ See, e.g., *United States v. Scheffer*, 523 U.S. 303, 312-13 (1998) ("A fundamental premise of our criminal trial system is that 'the jury is the lie detector.'" (quoting *United States v. Barnard*, 490 F.2d 907, 912 (9th Cir. 1973) (emphasis added by *Scheffer* Court))); George Fisher, *The Jury's Rise as Lie Detector*, 107 YALE L.J. 575 (1997).

³¹ The wrongful conviction of factually innocent defendants has recently been much in the news as a result of the numerous successes of the National Innocence Project and its many regional affiliates. See, e.g., Fernanda Santos, *Vindicated by DNA, but a Lost Man on the Outside*, N.Y. TIMES, Nov. 25, 2007, at 1; Abby Aguirre et al., *Exonerated, Freed, and What Happened Then*, N.Y. TIMES Interactive Feature (Nov. 25, 2007), http://www.nytimes.com/interactive/2007/11/25/nyregion/20071125_DNAL_FEATURE.html; see also *infra* note 195 (surveying popular media accounts of DNA exoneration). In addition, there is a growing scholarly literature on false conviction. See, e.g., BARRY SCHECK ET AL., *ACTUAL INNOCENCE: FIVE DAYS TO EXECUTION AND OTHER DISPATCHES FROM THE WRONGLY CONVICTED* (2000); Jules Epstein, *The Great Engine that Couldn't: Science, Mistaken Identifications, and the Limits of Cross-Examination*, 36 STETSON L. REV. 727 (2007); Keith A. Findley, *Innocents at Risk: Adversary Imbalance, Forensic Science, and the Search for Truth*, 38 SETON HALL L. REV. 893 (2008); Brandon L. Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55 (2008); Paul Gianelli, *Wrongful Convictions and Forensic Science: The Need to Regulate Crime Labs*, 86 N.C. L. REV. 163 (2007); Samuel R. Gross et al., *Exonerations in the United States, 1989 Through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523 (2005); D. Michael Risinger, *Innocents Convicted: An Empirically Justified Factual Wrongful Conviction Rate*, 97 J. CRIM. L. & CRIMINOLOGY 761 (2007); Michael J. Saks & D. Michael Risinger, *Baserates, the Presumption of Guilt, Admissibility Rulings, and Erroneous Convictions*, 2003 MICH. ST. L. REV. 1051. Whether such exonerations, including several of defendants previously sentenced to death, reveal serious systemic problems or only discreet and relatively rare instances of error is highly contested. See Morris B. Hoffman, *The Myth of Factual Innocence*, 82 CHI.-KENT L. REV. 663, 665 (2008) ("[T]he political subtext of the average innocence project is not that infrequent wrongful convictions must be detected and remedied, but rather that the system as a whole is profoundly unreliable and that the factual innocence revealed by the projects is just the tip of the iceberg of injustice."). Supreme Court Justice Antonin Scalia has written that the ultimate reversal of a wrongful conviction "demonstrates not the failure of the system but its success." *Kansas v. Marsh*, 548 U.S. 163, 193 (2005) (Scalia, J., concurring). It is probably impossible to reliably ascertain the true number or percentage of convicted defendants who are factually innocent. See Samuel R. Gross, *Convicting the Innocent*, 4 ANN. REV. L. & SOC. SCI. 173, 179 (2008) (describing the available data on false convictions, and noting that "once we move beyond murder and rape cases, we know very little about any aspect of false convictions") For a gripping firsthand account of an innocent man exonerated by DNA evidence after sixteen years in Georgia state prison, see CALVIN C. JOHNSON JR. WITH GREG HAMPIKIAN, *EXIT TO FREEDOM* (2003).

sharper relief. Whether these less obvious functions might be better carried out in darkness than in light is a further question that complicates the analysis.³²

Following this Introduction, Part I presents a summary of recent neuroscientific advances in lie detection and other “mind reading” research. Without attempting to evaluate all of the strengths and weaknesses of the existing studies,³³ this Part seeks primarily to demonstrate the very rapid rate at which the research has progressed over the past few years, and to suggest that there is a real possibility that the technology will ultimately advance to a stage of forensic viability. Should this happen, courts will be faced with difficult choices about admitting evidence generated by such technologies.³⁴

³² As discussed *infra* in Part III, many courts treat jury nullification with a kind of “don’t ask, don’t tell” attitude: the power to nullify exists, and occasionally it may be exercised in a laudatory manner, but jurors certainly should not be informed of this power and indeed should be told that they are bound to follow the law as set out by the court. See *infra* notes 236–49 and accompanying text.

³³ Others have amply covered this ground. See, e.g., ALBERT VRIJ, DETECTING LIES AND DECEIT: PITFALLS AND OPPORTUNITIES 365–87 (2d ed. 2008); Henry T. Greely & Judy Illes, *Neuroscience-Based Lie Detection: The Urgent Need for Regulation*, 33 AM. J.L. & MED. 377 (2007); Daniel D. Langleben, *Detection of Deception with fMRI: Are We There Yet?*, 13 J. LEGAL & CRIM. PSYCH. 1 (2008) [hereinafter Langleben, *Are We There*]; Sean A. Spence, *Playing Devil’s Advocate: The Case Against fMRI Lie Detection*, 13 J. LEGAL & CRIM. PSYCH. 11 (2008) [hereinafter Spence, *Devil’s Advocate*]; Paul Root Wolpe et al., *Emerging Neurotechnologies for Lie-Detection: Promises and Perils*, 5 AM. J. BIOETHICS 39 (2005). Generally, the methodological and other limitations advanced by critics have been acknowledged by the authors in the papers themselves and have been identified as areas for further study, some of which has already begun. Limitations include problems of generalizability from the populations studied, which have largely been young, healthy, drug-free, right-handed men with no brain abnormalities or personality disorders; the current need for compliant subjects willing to lie still inside the magnet; the “low-stakes,” investigator-sanctioned and impersonal nature of the deceptive behavior studied; and the averaging of results across groups and across trials rather than by individual subject and single trial. In addition to these methodological challenges, ethical issues raised by the new neuroscience have been the subject of recent discussion. See Henry T. Greely, *The Social Effects of Advances in Neuroscience: Legal Problems, Legal Perspectives*, in NEUROETHICS: DEFINING THE ISSUES IN THEORY, PRACTICE, AND POLICY (Judy Illes ed., 2006); Raymond De Vries, *Who Will Guard the Guardians of Neuroscience? Firing the Neuroethical Imagination*, 8 EUR. MOLECULAR BIOLOGY ORG. REP. S65 (2007); Martha J. Farah, *Emerging Ethical Issues in Neuroscience*, 5 NATURE NEUROSCI. 1123 (2002); *Neuroethics Needed*, 441 NATURE 907 (2006); Eric Racine et al., *fMRI in the Public Eye*, 6 NATURE REV. NEUROSCI. 159 (2005); Sean A. Spence & Catherine J. Kaylor-Hughes, *Looking for Truth and Finding Lies: The Prospects for a Nascent Neuroimaging of Deception*, 14 NEUROCASE 68, 78 (2008) [hereinafter Spence & Kaylor-Hughes, *Looking for Truth and Finding Lies*] (“The . . . question (‘should we do it?’) is one for society at large [I]f one is concerned with medical ethics and the proper treatment of human subjects, then it is difficult to foresee a situation in which it would be justifiable to scan people against their will.”).

³⁴ In the case of polygraph, which has thus far been the only even arguably viable forensic lie-detection technology, courts inclined to exclude the evidence (that is, most courts) have had at their disposal a relatively straightforward reliability rationale. See Jeffrey Bellin, *The Significance (If Any) for the Federal Criminal Justice System of Advances in Lie Detector Technology*, 80 TEMP. L. REV. 711, 711–12 (2007) (noting that “[f]or decades, so called ‘lie detector’ evidence has been barred in federal and state courts on the ground that traditional lie detector technology is unreliable, amounting to little more than junk science,” and that “courts have rarely gone beyond the reliability determination to evaluate whether, once the science of lie detection

Accordingly, the remainder of this Article proceeds in the nature of a thought experiment.³⁵ Assuming that there may come a time when a machine becomes far superior to a jury at detecting lies,³⁶ courts and scholars will be faced with

improves, lie detector evidence will be admissible”); *see also* Paul C. Gianelli, *Polygraph Evidence: Post-Daubert*, 49 HASTINGS L.J. 895 (1998) (surveying case law). The few courts that have squarely confronted the question of admissibility of advance neuroimaging or other brain-based lie detection evidence have excluded it as unreliable. *See* Harrington v. State, 659 N.W.2d 509, 516 n.6 (Iowa 2003); Slaughter v. State, 105 P.3d 832, 835–36 (Okla. Crim. App. 2005). In other contexts, however, admissibility of neuroscientific evidence is more common. *See generally* Khoshbin & Khoshbin, *supra* note 18, at 181–86 (noting that “the number of cases allowing defendants to present neuroimaging evidence is growing” and discussing several examples); Moriarty, *supra* note 17, at 29.

³⁵ As this is a thought experiment, I have put aside issues of (economic) cost, as well as certain inconvenient technical details. (For example, in my ideal, imaginary courtroom the lie detector would be portable—much like a headband—and would be used *during* witness testimony so that the truth or falsity of the testimony would be immediately apparent.) *See* Sharon Begley, *Mind Reading Is Now Possible*, NEWSWEEK, Jan. 21, 2008, at 22, available at <http://www.newsweek.com/id/91688> (“Nothing in physics rules out remote detection of brain activity. In fact, says law professor Hank Greely of Stanford, an infrared device under development might read thoughts using little more than a headband.”); Greely & Illes, *supra* note 33, at 388–89 (describing technique of Near-infrared Spectroscopy (NIRS)). In the real world, MRI remains expensive, running in the neighborhood of \$700 for a single scan. *See* Reed Abelson, *An M.R.I. Machine for Every Doctor? Someone Has to Pay*, N.Y. TIMES, Mar. 13, 2004, at A1. And portability and subject movement are not currently feasible. *See infra* note 63 (citing barriers to the forensic use of the technology as it currently stands). *But see* Yoko Hoshi et al., *Movable Cognitive Studies with a Portable, Telemetric Near-Infrared Spectroscopy System*, 13 NEUROIMAGE S18 (2001) (reporting an experiment with child subjects using a non-invasive, wireless neuroimaging system that does not require restriction of movement); K. Luan Phan et al., *Neural Correlates of Telling Lies: A Functional Magnetic Resonance Imaging Study at 4 Tesla*, 12 ACAD. RADIOL. 164 (2005) (using a novel real-time fMRI technology that allows researchers to obtain brain-activation results while the subject is inside the scanner, rather than having to analyze the data and generate results after the experiment).

³⁶ Humans in general are not very good at detecting deception, and furthermore are much worse than we think we are. *See* VRII, *supra* note 33, at 152, 156 (summarizing extensive psychological literature demonstrating that people generally perform only slightly better than chance at classifying lies and truth in both strangers and intimates); *id.* at 2 (“People tend to overestimate their own ability to detect lies . . .” (citing Eitan Elaad, *Effects of Feedback on the Overestimated Capacity to Detect Lies and the Underestimated Ability to Tell Lies*, 17 APPLIED COGNITIVE PSYCH. 349, 350 (2003))); Saul M. Kassir, *Human Judges of Truth, Deception, and Credibility: Confident but Erroneous*, 23 CARDOZO L. REV. 809 (2002). On the question whether certain individuals (such as judges or law enforcement officials), by virtue of experience, training, or inherent ability, are able to detect lies at levels significantly better than chance, compare Charles F. Bond Jr. & Bella M. DePaulo, *Individual Differences in Judging Deception: Accuracy and Bias*, 134 PSYCH. BULL. 477 (2008) (in a meta-analysis of research on individual differences in detecting deception, finding no evidence to support claim that certain classes of individuals are “truth wizards”), with Maureen O’Sullivan, *Home Runs and Humbugs: Comment on Bond and DePaulo (2008)*, 134 PSYCH. BULL. 493 (2008) (critiquing the statistical methodology of the meta-analysis and reaffirming previous findings of Eckman, O’Sullivan, and colleagues on the existence of “truth wizards”). Legal scholars have examined the social science research and have argued that the system’s reliance on juries to detect lies through demeanor evidence is largely misplaced. *See* Jeremy A. Blumenthal, *A Wipe of the Hands, A Lick of the Lips: The Validity of Demeanor Evidence in Assessing Witness Credibility*, 72 NEB. L. REV. 1157, 1159 (1993) (“The studies establish that typical subjects are unable to use the ‘manner and conduct’ of a speaker to successfully detect deceptive information on any reliable basis.”); Fisher, *supra* note 30, at 707 (“We could perhaps regard the wonderful convenience of jury

several thorny questions. First is whether such evidence will—or should—be excluded even assuming its accuracy and reliability.³⁷ Courts have traditionally been exceptionally wary of evidence that touches upon witness credibility; it is likely that this resistance will also play out in the context of brain-based lie detection evidence. Part II examines the justice system's proclivity to keep the lid of the jury's black box tightly sealed. It recounts the longstanding suspicion of expert evidence concerning witness credibility and the rules that guard jury secrecy, and concludes that these have as much to do with protecting the inviolability of the black box of the jury as with guarding against unreliable evidence. Part II goes on to discuss several developments that have increasingly put pressure upon the jury as sole fact finder and lie detector, such that the availability of a more reliable lie detector would ultimately be difficult to dismiss. Assuming such a development, this Part addresses the question whether, in cases in which credibility of witnesses is central to the outcome, the jury would remain a viable institution despite the existence of a highly accurate lie detection machine. It analyzes the arguments of scholars who have asserted that juries would continue to perform crucial functions, and suggests that these remaining jury functions would be more in the nature of the jury's legislative than adjudicative role.³⁸

lie detecting with more equanimity if there were any sound evidence that juries are *good* at this task. But most of the evidence we have suggests that juries have no particular talent for spotting lies. Not only do experimental subjects rarely perform much better than chance at distinguishing truth from falsehood, but they *think* they are better lie detectors than they are." (citations omitted); Olin Guy Wellborn III, *Demeanor*, 76 CORNELL L. REV. 1075, 1075 (1991) ("According to the empirical evidence, ordinary people cannot make effective use of demeanor in deciding whether to believe a witness. On the contrary, there is some evidence that the observation of demeanor diminishes rather than enhances the accuracy of credibility judgments."). Some studies have suggested that accuracy increases in the group context and when the individuals in a group are able to deliberate, but the question is far from settled. See generally Cass R. Sunstein & Reid Hastie, *Four Failures of Deliberating Groups* (Univ. of Chicago Pub. Law & Legal Theory Working Paper No. 215, 2008), http://ssrn.com/abstract_id=1121400 (noting that many deliberating groups do worse than their members' pre-deliberation judgments). In any event, even a generous view of the data suggests that as reliable lie detectors go, the jury probably does not present a very high bar.

³⁷ One scholar, for example, has argued that the primary obstacle to admission of such evidence would be the rule against hearsay, but that the evidence could be admitted under the residual exception contained in Federal Rule of Evidence 807. See Bellin, *supra* note 34. But see Edward J. Imwinkelried & James R. McCall, *Issues Once Moot: The Other Evidentiary Objections to the Admission of Exculpatory Polygraph Examinations*, 32 WAKE FOREST L. REV. 1045 (1997) (arguing that the rule against hearsay would not bar polygraph evidence that satisfies the *Daubert* reliability test, nor would Rules 402 or 608). Other objections to reliable fMRI lie detection evidence center less on the subtleties of the rules of evidence and more on issues of privacy, concerns about criminal defendants' constitutional rights, and the desire to preserve the traditional role of the jury. See discussion *infra* note 213 and accompanying text.

³⁸ These terms are adapted from the classic scholarship of Professor Kenneth Culp Davis on judicial notice and administrative law. See Kenneth Culp Davis, *Judicial Notice*, 1969 L. & SOC. ORDER 513 [hereinafter Davis, *Judicial Notice II*]; Kenneth Culp Davis, *Judicial Notice*, 55 COLUM. L. REV. 945 (1955)

Turning to focus on this quasi-legislative role of the jury, Part III then takes up the issue of jury nullification as a window into this more hidden purpose of the jury system. Jury decisions that go beyond the fact-finding role to contradict the legal instructions of the court are typically viewed with acute ambivalence—if not outright hostility—by most courts and some scholars.³⁹ Such jury behavior, which goes by the somewhat misleading label “jury nullification,”⁴⁰ has been described by the Supreme Court as a power, but not a right.⁴¹ Historical instances of nullification by juries are held up as the best of

[hereinafter Davis, *Judicial Notice* I]. They correspond roughly to the jury’s fact-finding (adjudicative) and law making (legislative) functions. See *infra* text accompanying note 223; see also Kenneth Culp Davis, *An Approach to Problems of Evidence in the Administrative Process*, 55 HARV. L. REV. 364, 404–07 (1942) [hereinafter Davis, *An Approach to Problems*]. Indeed, the jury has been called a “little parliament.” See VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 21 (quoting PATRICK DEVLIN, *TRIAL BY JURY* 164–65 (1956)); see also D. Michael Risinger, *Unsafe Verdicts: The Need for Reformed Standards for the Trial and Review of Factual Innocence Claims*, 41 HOUS. L. REV. 1281, 1291 (2005) (noting that where a jury decides such issues as negligence, which necessarily entail normative judgment overlaid upon findings of historical fact, “[i]t is like a particularized legislature for the particular circumstances of the case”).

³⁹ See *infra* notes 237–52 and accompanying text. Courts are almost uniformly hostile to nullification, though they sometimes pay lip service to historical instances of “good” nullification, such as the refusal of many northern antebellum juries to convict under the Fugitive Slave Act. See, e.g., *United States v. Dougherty*, 473 F.2d 1113, 1130 (D.C. Cir. 1972) (holding that the defendants had no right to have the jury instructed on its power to nullify, while acknowledging that “[t]he pages of history shine on instances of the jury’s exercise of its prerogative to disregard uncontradicted evidence and instructions of the judge”). Scholars present more varied views, with many supporting the criminal jury’s role as the “conscience of the community,” but others are sharply critical of nullification, seeing it as an example of anarchy and lack of accountability. Compare Clay S. Conrad, *Scapegoating the Jury*, 7 CORNELL J.L. & PUB. POL’Y 7 (1997), with Judge Lawrence W. Crispo et al., *Jury Nullification: Law Versus Anarchy*, 31 LOY. L.A. L. REV. 1 (1997).

⁴⁰ The label is not only imprecise, but also has a derogatory connotation. See *United States v. Polizzi*, 549 F. Supp. 2d 308 (E.D.N.Y. 2008) (“Placing the pejorative characterization of ‘nullification’ on the jury’s Sixth Amendment power [to convict of a lesser crime in order to reduce a defendant’s sentence] does not define it out of existence.”); GEORGE P. FLETCHER, *A CRIME OF SELF-DEFENSE: BERNHARD GOETZ AND THE LAW ON TRIAL* 154 (1988) (arguing that the term “jury nullification” is misleading because it connotes “an act of disrespect toward the law”); CLAY S. CONRAD, *JURY NULLIFICATION: THE EVOLUTION OF A DOCTRINE* 6 (1998) (“It is both derisive and deceptive to refer to the discretionary powers of the jury as ‘jury nullification.’ It is derisive because it gives a very negative description of what the jury does, and it assumes that the jury is active outside their legal powers [It is deceptive because] it is the jury which nullifies that particular application of the statute, and not the jury which is nullified.”). Conrad, who has argued in favor of the criminal jury’s right and obligation to judge the law, has proposed the term “jury independence.” See *id.*

⁴¹ See *Sparf & Hansen v. United States*, 156 U.S. 51, 80 (1895) (holding that trial judge’s refusal to instruct the jury that it could return a verdict on lesser charge of manslaughter was not error where such verdict was not supported by the evidence, and approvingly describing Massachusetts Supreme Court Justice Shaw’s observation that “though the jury ha[d] the power they had not the right to decide, that is, to adjudicate on both law and evidence”); see also *United States v. Washington*, 705 F.2d 489, 494 (D.C. Cir. 1983) (per curiam) (“A jury has no more ‘right’ to find a ‘guilty’ defendant ‘not guilty’ than it has to find a ‘not guilty’ defendant ‘guilty,’ and the fact that the former cannot be corrected by a court, while the latter can be, does not create a right out of the power to misapply the law.”). But see CONRAD, *supra* note 40, at 301 (arguing that “[t]he

what the jury system represents,⁴² and as the worst.⁴³ This Part proposes that attitudes toward jury nullification are an extreme example of a much more pervasive distrust of juries⁴⁴ and a profound ambivalence about their role in our justice system. Because the extent of nullification has thus far largely been hidden inside the black box of the jury room, much of this ambivalence often simmers beneath the surface.⁴⁵ However, the more that newly developed scientific and forensic technologies remove genuine doubt about historical facts at trial, the more obvious is the jury's ability to engage in nullification. This Part considers the implications, in terms of legitimacy, institutional function, and the constitutional right to trial by jury, of such a development.

distinction between jury 'rights' and jury 'powers' is nonsensical, and should be discarded" because "[a] legal power that can be exercised with legal impunity is a legal right").

⁴² The 1670 acquittal of William Penn and the refusal to convict John Peter Zenger of seditious libel are frequently cited examples. See VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 225.

⁴³ Southern juries' refusals to convict those charged with lynchings and other racial violence throughout much of the twentieth century are the most common historical examples. See *id.* ("If Northern juries of the 1850s acquitted the courageous people who harbored slaves in defiance of the Fugitive Slave Act, Southern juries a hundred years later acquitted people who had beaten and killed the descendants of those slaves as they attempted to assert their legal rights."); cf. *United States v. Navarro-Vargas*, 408 F.3d 1184, 1199 (9th Cir. 2005) (en banc) (noting in the context of the grand jury's power to nullify that "[w]hile we celebrate grand jury independence in defense of the First Amendment in the case of Peter Zenger and those accused of violating the Alien and Sedition Acts, and we praise grand jury resistance to the morally-obnoxious fugitive slave laws, we must acknowledge as well that grand juries have also refused to enforce lawful and wise legislation, including some of the most important legislation in American history: the Reconstruction laws implementing the Thirteenth, Fourteenth, and Fifteenth Amendments. Grand jury independence, evidently, has historically served causes both good and ill.").

⁴⁴ See Julie A. Seaman, *Triangulating Testimonial Hearsay: The Constitutional Boundaries of Expert Opinion Testimony*, 96 GEO. L.J. 827, 869–76 (2008) (discussing distrust of juries in the context of expert witness testimony); see also *United States v. Dougherty*, 473 F.2d 1113, 1142 (D.C. Cir. 1972) (Bazelon, J., concurring in part and dissenting in part) ("Trust in the jury is, after all, one of the cornerstones of our entire criminal jurisprudence, and if that trust is without foundation we must re-examine a great deal more than just the nullification doctrine.").

⁴⁵ It tends to resurface when a jury in a notorious case delivers a verdict that the public and the press believe is at odds with the evidence or during times of social tension when significant numbers of citizens disagree about particular laws or policies. E.g., Paula L. Hannaford-Agor & Valerie P. Hans, *Nullification at Work? A Glimpse from the National Center for State Courts Study of Hung Juries*, 78 CHL-KENT L. REV. 1249, 1251 (2003) ("Several high-profile trials in recent years have contributed to the perception that jury nullification is increasing."); Alan W. Schefflin & Jon M. Van Dyke, *Merciful Juries: The Resilience of Jury Nullification*, 48 WASH. & LEE L. REV. 165, 175 (1991) (noting that the issue of jury nullification "lay dormant for most of [the twentieth] century until resurrected in the 1960s as part of the defense strategy in anti-Vietnam War demonstration trials").

I. INSIDE THE FIRST BLACK BOX: WATCHING THE BRAIN THINK

Over the past decade or so, scientists have been busy “mapping” the brain,⁴⁶ primarily using a technique known as functional magnetic resonance imaging (fMRI).⁴⁷ In contrast to structural magnetic resonance imaging

⁴⁶ See generally Arthur W. Toga et al., *Towards Multimodal Atlases of the Human Brain*, 7 NATURE REV. NEUROSCI. 952 (2006). This mapping process is accomplished by investigating the neural correlates of various responses or behaviors. As the term suggests, such studies reveal correlations between a mental or physical activity and particular brain structures; as the research accumulates, scientists have begun to hypothesize about the functions of various brain structures. In addition, functions may be mapped by studying individuals with localized brain injuries to determine the deficits caused by such injuries. See, e.g., THE NEUROPSYCHOLOGY HANDBOOK 4–6 (Arthur MacNeill Horton, Jr. & Danny Wedding eds., 2008) (describing early discoveries of localized brain function through inspection of patients’ brains after they had died). Finally, new technologies allow researchers to stimulate parts of the brain temporarily and thereby illuminate the functions of those areas. See Alberto Priori et al., *Lie-Specific Involvement of Dorsolateral Prefrontal Cortex in Deception*, 18 CEREBRAL CORTEX 451, 451 (2008) (employing transcranial direct current stimulation (tDCS), “a noninvasive technique that elicits functional changes in the human brain, without needing a direct access to the neural tissue”). Though there is still much to be learned, it is widely accepted for example that so-called “executive” functions are centered in the pre-frontal cortex, that the amygdala is involved in the experience of emotions such as fear and disgust, and that “the left hemisphere has superior language and arithmetic skills, whereas the right hemisphere has better spacial skills.” Chris Rorden & Hans-Otto Karnath, *Using Human Brain Lesions to Infer Function: A Relic from a Past Era in the fMRI Age?*, 5 NATURE REV. NEUROSCI. 812 (2004); see also, e.g., Andrea C. Pierno et al., *Observing Social Interactions: The Effect of Gaze*, 3 SOC. NEUROSCI. 51, 51 (2008) (medial prefrontal cortex involved in social cognition); Robert M. Sapolsky, *The Frontal Cortex and the Criminal Justice System*, in LAW & THE BRAIN (Semir Zeki & Oliver R. Goodenough eds., 2006) (summarizing accepted functions of pre-frontal cortex). However, there is concern that fMRI studies can lead to overblown claims, particularly in the popular press. See, e.g., *Growing Pains for fMRI*, 320 SCIENCE 1412, 1412 (2008) (discussing criticism by a group of neuroscientists of a column in the *New York Times* that had reported a study of voter attitudes toward presidential candidates drawn from fMRI images of undecided voters’ brain activity, and quoting one neuroscientist as saying that the study’s conclusion “was really closer to astrology than it was to real science,” and that it “epitomized everything that a lot of us feel is wrong about where certain parts of the field are going, which is throw someone in a scanner and tell a story about it”); Weisberg et al., *supra* note 18, at 470 (discussing a study in which both lay persons and experts judged bad explanations as good where neuroscience language was added to the explanation). These attempts to situate particular human activities in specific parts of the brain have been the subject of some significant criticism and debate. See, e.g., Goodenough, *supra* note 4, at 432–35 (discussing contrast between Cartesian model of unified mind and neuroscience model of modular mind). The debate mirrors one that surrounds the field of evolutionary psychology and critique of genetic reductionism. See Julie A. Seaman, *Form and (Dys)Function in Sexual Harassment Law: Biology, Culture, and the Spandrels of Title VII*, 37 ARIZ. ST. L.J. 321, 345–46 (2005).

⁴⁷ Functional magnetic resonance imaging is one of a number of brain-imaging techniques. Others include Positron Emission Tomography (PET), single photon emission computed tomography (SPECT), electroencephalography (EEG), and magnetoencephalography (MEG). See Society for Neuroscience, Brain Briefings (May 1996), http://www.sfn.org/index.cfm?pagename=brainBriefings_brainImaging; see also R. Christopher deCharms, *Applications of Real-Time fMRI*, 8 NATURE REV. NEUROSCI. 720 (2008). Because of its relative safety, non-invasiveness, and ability to pinpoint specific brain areas, research using fMRI outpaces that of other technologies. As a rough measure of the vibrancy of this area of research, there have been some 19,000 fMRI peer-reviewed papers published since the technique was first developed. “Given that the first

(fMRI), which generates a very detailed and high-resolution but static image of the brain,⁴⁸ a functional MRI generates images over time as the subject is *doing* something—for example, looking at a photograph of a face,⁴⁹ listening to music,⁵⁰ making an economic decision,⁵¹ thinking about a particular object,⁵² or telling a lie.⁵³ In this way, researchers attempt to discover which parts of the brain are more actively working while the subject is engaged in a given cognitive or motor activity or is experiencing a particular emotional or mental state.⁵⁴ These studies have revealed a wealth of tantalizing information about the inner workings of the human mind, with implications for a number of law-relevant areas. Scholars working at the intersection of law and neuroscience⁵⁵ have begun to explore the potential impact of neuroscientific

fMRI study without exogenous contrast agents was published in 1991, this corresponds to approximately 1,100 papers per year . . . [though the] average obscures the actual rate of publications, as in 1992 there were four publications in total, increasing to about eight per day by 2007.” Nikos K. Logothetis, *What We Can Do and What We Cannot Do with fMRI*, 453 *NATURE* 869, 869 (2008); see M.J. Friedrich, *Neuroscience Becomes Image Conscious as Brain Scans Raise Ethical Issues*, 294 *JAMA* 781 (2005) (noting explosion in use of fMRI since it first appeared in the early 1990s). There are now a number of peer-reviewed scientific journals devoted exclusively to brain-scanning papers, for example *NeuroImage* and *Human Brain Mapping*.

⁴⁸ Structural MRI is used for medical diagnosis and treatment, as well as in basic research. Logothetis, *supra* note 47, at 869.

⁴⁹ See Steven M. Platak et al., *Neural Substrates for Functionally Discriminating Self-Face from Personally Familiar Faces*, 27 *HUM. BRAIN MAPPING* 91 (2006) (finding that brain activation in response to a subject’s own face was distinct from that for a familiar face, and that brain activation for familiar faces was distinct from unknown faces).

⁵⁰ See Stefan Koelsch et al., *Adults and Children Processing Music: An fMRI Study*, 25 *NEUROIMAGE* 1068 (2005) (investigating and comparing neural activations in response to music in adults with no musical training, adults with musical training, and children with a range of musical training).

⁵¹ See, e.g., Eric Maskin, *Can Neural Data Improve Economics?*, 321 *SCIENCE* 1788 (2008).

⁵² See Shinkareva et al., *supra* note 16 (finding distinct neural activations where subjects viewed line drawings of various objects drawn from different categories, such as tools and dwellings).

⁵³ See studies discussed *infra* notes 62–124 and accompanying text.

⁵⁴ With the technique known as Blood Oxygenated Level Dependent (BOLD) fMRI, “the local change in the concentration of oxygenated hemoglobin in the brain is used as an indicator of neuronal activity.” Wolpe et al., *supra* note 33, at 43. Though this assumption that increased blood flow indicates increased neural activity is widespread and underlies the conclusions drawn from BOLD fMRI studies, the connection between blood flow and neuronal activity is not well-understood. See John-Dylan Haynes & Geraint Rees, *Decoding Mental States from Brain Activity in Humans*, 7 *NATURE REV. NEUROSCI.* 523, 529 (2006); *Blood Flow May Be Key Player in Neural Processing*, *SCIENTIFIC AM.*, Jan. 24, 2008, available at <http://www.sciam.com/article.cfm?id=blood-flow-may-be-key-player> (detailing theory that blood flow is more intimately involved in neural activity than previously thought).

⁵⁵ The field has been dubbed “neurolaw.” See Jeffrey Rosen, *The Brain on the Stand*, *N.Y. TIMES*, Mar. 11, 2007 (Magazine), at 50 (“[T]he influence of what some call neurolaw is clearly growing.”). It takes its place among other “neuroscience-and” fields such as neuroeconomics, neuroethics, neuropsychology, neurobiology, and neurolinguistics.

discoveries upon criminal responsibility and punishment,⁵⁶ economic incentives and trust,⁵⁷ property,⁵⁸ and discrimination,⁵⁹ among others.⁶⁰

A. *The Brain on Lies—An Overview*

Research into the neural correlates of deception is among the more active—and promising—subjects of brain imaging with potential legal applications.⁶¹ Since 2001, a number of brain-imaging studies have compared neural activity during truth-telling with that during deception.⁶² Within this

⁵⁶ See Dean Mobbs et al., *Law, Responsibility, and the Brain*, 5 PLOS BIOLOGY 693 (2007); Richard E. Redding, *The Brain-Disordered Defendant: Neuroscience and Legal Insanity in the Twenty-First Century*, 56 AM. U. L. REV. 51 (2006); Amanda C. Pustilnik, *Violence on the Brain: A Critique of Neuroscience in Criminal Law* (Harvard Law Sch. Faculty Scholarship Series, Paper No. 14, 2008), available at <http://lsr.nellco.org/harvard/faculty/papers/14/>.

⁵⁷ See Terrence Chorvat & Kevin McCabe, *The Brain and the Law*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y LONDON B: BIOLOGICAL SCI. 1727 (2004).

⁵⁸ See Jeffrey Evans Stake, *The Property 'Instinct'*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y LONDON B: BIOLOGICAL SCI. 1763 (2004) (arguing that property-related behaviors can be explained by empirical studies and evolutionary theory, and that fundamental principles of property are encoded in the human brain).

⁵⁹ See, e.g., Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489 (2005); see also Hart et al., *supra* note 11, at 2351; Phelps et al., *supra* note 11, at 729; J.A. Richeson et al., *An fMRI Investigation of the Impact of Interracial Contact on Executive Function*, 6 NATURE NEUROSCI. 1323 (2003).

⁶⁰ As with many past technological innovations that have made their way into the courtroom, the new neuroscience has been met with concern that its claims are overblown, see Stephen Morse, *Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note*, 3 OHIO ST. J. CRIM. L. 397 (2006), or that it will unduly sway gullible juries, see Khoshbin & Khoshbin, *supra* note 18, at 182–83. Professor Michael Pardo, addressing this concern, argues that neuroscience evidence, “[i]f properly understood, . . . may be properly assimilated into legal practices without undermining those practice or usurping powers of judges and juries.” Pardo, *supra* note 27, at 312 (noting the similar concerns raised about photographic evidence and DNA evidence when these were in their relative infancy). Professor Jennifer Mnookin has detailed the early response to photographic evidence, noting the argument made at the time that “the photograph’s reputation for truthfulness . . . risked making it ‘a most dangerous perjurer’” and “dangerous evidence indeed.” Jennifer L. Mnookin, *The Image of Truth: Photographic Evidence and the Power of Analogy*, 10 YALE J.L. & HUMAN. 1, 26 (1998).

⁶¹ I mean to roughly distinguish here between research that is of potential theoretical interest in shaping or critiquing legal rules, on the one hand, and research that might lead to courtroom or other direct applications in gathering evidence or resolving legal disputes.

⁶² A recent review article counts sixteen peer-reviewed data papers as comprising the fMRI literature on deception, including one in press at the time of the review that has since been published. See Spence, *Devil’s Advocate*, *supra* note 33, at 13. Since that review was written, three additional studies have become available. See Jeffrey N. Browndyke et al., *Neuroanatomical Correlates of Malingered Memory Impairment: Event-Related fMRI of Deception on a Recognition Memory Task*, 22 BRAIN INJURY 481 (2008) (finding distinctive neural activation in two types of malingering activity: false omission and false inclusion); J.G. Hakun et al., *Towards Clinical Trials of Lie Detection with fMRI*, 3 SOC. NEUROSCI. (forthcoming 2008); Sean A. Spence et al., *Speaking of Secrets and Lies: The Contribution of Ventrolateral Prefrontal Cortex to Vocal Deception*, 40 NEUROIMAGE 1411 (forthcoming 2008) [hereinafter Spence et al., *Speaking of Secrets*]; see also George T. Monteleone et al., *Detection of Deception Using fMRI: Better than Chance, but Well Below Perfection*, 3 SOC. NEUROSCI. (forthcoming 2008) (reanalyzing data from a previous study); Sam Harris, *Functional*

relatively brief time frame, great strides have been made, both in refining the experimental methodologies and in enhancing the technological and statistical methods used to generate the results. In this Part, I review the existing literature and describe these rapid advances in the field to demonstrate the remarkable pace at which it continues to progress. Although certainly the technology is not yet at a stage at which it can be applied for forensic purposes, if the research continues apace, applied uses are likely to become a real possibility.⁶³

Though specific results thus far have been somewhat varied—probably due to differences in experimental design and scanning parameters among the studies⁶⁴—some preliminary conclusions may be drawn:

[M]ost studies have reported that areas of prefrontal cortex exhibit greater activation during the lie response than telling the truth; most have reported no areas where activity is greater during truth telling than lying; and where investigators have measured response times (RTs), they have mostly reported longer RTs during deception than truth telling. All these findings are consistent with one central

Neuroimaging of Belief, Disbelief, and Uncertainty, 63 ANN. NEUROL. 141 (2008) (finding that the cognitive states of belief, disbelief, and uncertainty had distinct neural correlates). In addition, several published studies have employed EEG, PET, TMS, and MEG technologies to investigate the neural correlates of deception or distinctive brain-wave patterns correlated with a recognition response (so-called “brain-fingerprinting”). See, e.g., Karen J. Kelly et al., *The Effect of Deception on Motor Cortex Excitability*, 3 SOC. NEUROSCI. (forthcoming 2008) (using transcranial magnetic stimulation and finding hemispheric differences in motor cortex excitability where subjects answered deceptively about their favored sports team); Y.L. Lo et al., *Increased Cortical Excitability in Human Deception*, 14 NEUROREPORT 1021 (2003) (employing TMS and finding increased cortical excitability in deceptive as compared to truthful responses to a series of questions); Ming Ann Lui et al., *Thirty-Site P300 Scalp Distribution, Amplitude Variance Across Sites, and Amplitude in Detection of Deceptive Concealment of Multiple Guilty Items*, 3 SOC. NEUROSCI. (forthcoming 2008) (using scalp distribution approach to obtain both group and individualized distinctions between deception and truth); Ray Johnson Jr. et al., *The Self in Conflict: The Role of Executive Processes During Truthful and Deceptive Responses About Attitudes*, 39 NEUROIMAGE 469 (2008) (using EEG and finding differences in brain activation and magnitude of response for self-referential lies and truthful responses).

⁶³ Professors Sean Spence and Catherine Kaylor-Hughes, leading researchers in the area, acknowledge that there are many limitations to the current body of work. However, they view the technical problems involved with developing forensic uses for fMRI lie detection as “complex” but still “tractable.” See Spence & Kaylor-Hughes, *Looking for Truth and Finding Lies*, *supra* note 33. Indeed, their group has recently completed a forensic imaging study, in which they examined a woman who had been convicted of child abuse and was said to be suffering from Munchausen’s syndrome by proxy. The results were consistent with innocence. See Spence et al., *Munchausen’s Syndrome*, *supra* note 17.

⁶⁴ Jennifer Maria Nuñez et al., *Intentional False Responding Shares Neural Substrates with Response Conflict and Cognitive Control*, 25 NEUROIMAGE 267, 267 (2005) (describing prior studies and suggesting that the variability seen in activation of different brain regions across studies “may be due in part to the large variability among experimental designs as well as their broad and behaviorally complex definitions of a deceptive act”).

hypothesis that lying “behaves” like an executive task, requiring the contribution of “higher” centres, and incurring a cognitive cost in terms of processing time.⁶⁵

B. *The Brain-Imaging Studies*

In the earliest fMRI deception research, scientists sought to build on the findings of polygraph and scalp-recorded event-related potential (ERP) research that suggested there were “cognitive differences between lying and telling the truth [that] could be associated with changes in other correlates of brain activity.”⁶⁶ Because polygraphs and ERPs are relatively indirect and rough measures, respectively,⁶⁷ researchers hoped to use fMRI to pinpoint those brain structures involved in deception. In addition, fMRI held the promise of overcoming certain limitations of the existing deception-detection technologies. First, because polygraphy essentially measures anxiety, and because “deception-induced mood and somatic states may vary across individuals,” fMRI researchers hoped to isolate “a marker of deception independent of anxiety or guilt”⁶⁸ that would overcome this inherent limitation of polygraphy. Previous work on deception had suggested that intentional deception requires the cognitive process of inhibiting the truthful response regardless of feelings of guilt or anxiety;⁶⁹ because such inhibitory response

⁶⁵ Spence, *Devil’s Advocate*, *supra* note 33, at 22 (citation omitted); *see also* Langleben, *Are We There*, *supra* note 33, at 3 (stating that the fMRI lie detection research has “converged on a model of deception as a working memory-intensive task, mediated to a large extent by the profronto-parietal systems dedicated to behavioural control and attention”); Chris William Sanchirico, *Evidence, Procedure, and the Upside of Cognitive Error*, 57 STAN. L. REV. 291 (2004) (arguing that these features of deception and memory can be a boon to the legal system as a whole).

⁶⁶ D.D. Langleben et al., *Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study*, 15 NEUROIMAGE 727, 727 (2002) [hereinafter Langleben et al., *Brain Activity*].

⁶⁷ Polygraph is indirect—it measures peripheral nervous responses, such as sweating, heart rate, and respiration, rather than central nervous responses. *Id.* ERPs, which are measured by electroencephalography (EEG), are a direct measure of neural electrical activity in the brain but, because these electrical changes are measured from outside the scalp, “their source in the brain cannot be uniquely localized.” *Id.* The BOLD response used by fMRI is a measure of blood flow in the brain that is believed to directly correlate with neural activity; it therefore is a *more* direct signal of specific neural activity than polygraph, but one could argue it remains an indirect measure. *But see Blood Flow May Be Key Player in Neural Processing*, *supra* note 54 (noting that previously fMRI researchers had been slightly apologetic about the fact that they measured blood flow rather than neural signals themselves, but that new evidence suggests that blood flow may be directly involved in neural processing).

⁶⁸ Langleben et al., *Brain Activity*, *supra* note 66, at 728.

⁶⁹ *Id.* at 727; *see also* Sean A. Spence et al., *Behavioural and Functional Anatomical Correlates of Deception in Humans*, 12 NEUROREPORT 2849, 2849 (2001) [hereinafter Spence et al., *Anatomical Correlates*] (proposing that deception involves, among other cognitive processes, “the inhibition of pre-potent responses (truths)”).

had previously been correlated with specific areas of the prefrontal cortex,⁷⁰ the researchers hypothesized that deception under fMRI would differentially activate those areas. Indeed, this was the result.⁷¹ Second, because fMRI is a direct measure of brain activity rather than a measure of peripheral nervous response,⁷² researchers hoped that the procedure would be more resistant to countermeasures.⁷³ The existing body of published fMRI deception studies⁷⁴ can be roughly grouped into four clusters, which are described below.

⁷⁰ Spence et al., *Anatomical Correlates*, *supra* note 69, at 2852.

⁷¹ See Langleben et al., *Brain Activity*, *supra* note 66, at 727–28, 730; Spence et al., *Anatomical Correlates*, *supra* note 69, at 2851.

⁷² E.g., Daniel D. Langleben et al., *True Lies: Delusions and Lie-Detection Technology*, 34 J. PSYCHIATRY & L. 351, 357 (2006) [hereinafter Langleben et al., *True Lies*] (“Since deception is a cognitive phenomenon that takes place in the brain, the potential of the [polygraph measurement of blood pressure, skin conductance, heart rate, and breathing] in a lie-detection system is theoretically inferior to the more proximal, central nervous system (CNS) correlates of brain activity that could be obtained by EEG and fMRI.”).

⁷³ See Feroze B. Mohamed et al., *Brain Mapping of Deception and Truth Telling About an Ecologically Valid Situation: Functional MR Imaging and Polygraph Investigation—Initial Experience*, 238 RADIOLOGY 679, 685 (2006) (arguing that, in contrast to polygraph, “[i]t is likely that a subject cannot mask functional MR imaging brain activation patterns. . . . [T]he brain areas that are active during deception will always be active when the subject tells a lie [and] the same areas will always be inactive when the subject tells the truth”). The ability of subjects undergoing polygraph testing to influence the results through the use of countermeasures is a major issue in the debate over polygraph reliability. See, e.g., Charles Robert Honts et al., *Effects of Physical Countermeasures on the Physiological Detection of Deception*, 70 J. APPLIED PSYCHOL. 177 (1985); F. Andrew Kozel et al., *A Pilot Study of Functional Magnetic Resonance Imaging Brain Correlates of Deception in Healthy Young Men*, 16 J. NEUROPSYCHIATRY & CLINICAL NEUROSCI. 295, 296 (2004) [hereinafter Kozel et al., *Pilot Study*]; Ralf Mertens & John J.B. Allen, *The Role of Psychophysiology in Forensic Assessments: Deception Detection, ERPs, and Virtual Reality Mock Crime Scenarios*, 45 PSYCHOPHYSIOLOGY 286, 287 (2008). In the criminal context, it has generally been defendants who have sought to introduce polygraph evidence of innocence. See, e.g., *Banks v. United States*, 531 F.2d 1336 (5th Cir. 1976) (upholding denial of motion to vacate sentence despite appellant’s presentation of post-conviction polygraph evidence of his innocence); *Poindexter v. Booker*, 2007 WL 155667, at *4 (E.D. Mich. May 30, 2007) (noting petitioner’s earlier claim of innocence was verified by polygraph evidence, but granting petition for conditional writ of habeas corpus on other grounds); Robert N. Clinton, *The Right to Present a Defense: An Emergent Constitutional Guarantee in Criminal Trials*, 9 IND. L. REV. 711, 812 (1976) (“[T]he bulk of the criminal cases involving such evidence seems to involve cases in which test results of these kinds were offered by the accused.”); Debra J. Katz, *State v. Dean: A Compulsory Process Analysis of the Inadmissibility of Polygraph Evidence*, 1984 WIS. L. REV. 237, 246 (noting that polygraph evidence may be favorable and material to the defense). Much of the extreme judicial resistance to this evidence may thus be attributable to a fear that guilty defendants will manage to manipulate evidence to secure a wrongful acquittal. Cf. FED. R. EVID. 804(b)(3) (requiring corroboration of statements against interest in criminal cases where exculpatory hearsay statement is offered by the accused). There has thus far been only one study touching on the efficacy of countermeasures in fMRI deception testing. See G. Ganis et al., *Neural Correlates of Different Types of Deception: An fMRI Investigation*, 13 CEREBRAL CORTEX 830, 831 (2003) (comparing rehearsed to spontaneous lies). Research in the EEG/ERP area is a bit more advanced, though still preliminary. It generally suggests that accuracy of the results is affected when subjects are taught simple countermeasures. See Mertens & Allen, *supra* (finding guilty (deception) verdicts very likely to be accurate despite mental and physical countermeasures, but that innocent (truthful) verdicts were highly influenced by simple countermeasures and thus unlikely to be useful in practice); J. Peter Rosenfeld et al., *Simple, Effective Countermeasures to P300-Based Tests of Detection of*

1. “Guilty Knowledge” Playing Card Studies

The first group consists of studies which employ variations on a concealed-information playing-card protocol,⁷⁵ a type of “guilty knowledge” paradigm whereby the subject is instructed to deny something that the subject knows to be true.⁷⁶ Under this paradigm, subjects are given an envelope containing a

Concealed Information, 41 PSYCHOPHYSIOLOGY 205 (2004) (finding that countermeasures lowered detection rates dramatically); see also NAT’L RESEARCH COUNCIL, THE POLYGRAPH AND LIE DETECTION 175 (2003) (noting that “[s]ome believe that event-related potentials are less vulnerable to countermeasures than the polygraph,” but that “[t]he basic science, however, is unclear on whether or not people can learn to manipulate event-related potentials”); M. Sasaki et al., *Effects of a Mental Countermeasure on the Physiological Detection of Deception Using the Event-Related Brain Potentials*, 72 JAPANESE J. PSYCH. 322 (2001) (cited in Mertens & Allen, *supra*, at 287) (original in Japanese) (finding that mental countermeasures did not affect P300 amplitude or latency). But see Ming Lui & J. Peter Rosenfeld, *Detection of Deception About Multiple, Concealed, Mock Crime Items, Based on a Spatial-Temporal Analysis of ERP Amplitude and Scalp Distribution*, 45 PSYCHOPHYSIOLOGY 721 (2008) (employing scalp distribution measurement technique in order to reduce ability of subjects to employ countermeasures). Interestingly, it has been suggested that fMRI might actually be used to detect countermeasures in connection with other types of lie-detection techniques, because the use of countermeasures might be associated with specific “brain signatures.” See NAT’L RESEARCH COUNCIL, *supra*, at 174.

⁷⁴ A Japanese group led by Nobuhito Abe used PET to study the neural correlates of deception. See Nobuhito Abe et al., *Deceiving Others: Distinct Neural Responses of the Prefrontal Cortex and Amygdala in Simple Fabrication and Deception with Social Interactions*, 19 J. COGNITIVE NEUROSCI. 287 (2007). PET provides similar temporal and spatial resolution to fMRI, and for purposes of brain-mapping study of the neural correlates of deception it is comparable (though its forensic use is unfeasible because of the need to expose subjects to radioactive dyes). Therefore, this section references the Abe PET study where relevant.

⁷⁵ See C. Davatzikos et al., *Classifying Spatial Patterns of Brain Activity with Machine Learning Methods: Application to Lie Detection*, 28 NEUROIMAGE 663 (2005); Matthias Gamer et al., *Covariations Among fMRI, Skin Conductance, and Behavioral Data During Processing of Concealed Information*, 28 HUM. BRAIN MAPPING 1287 (2007); Langleben et al., *Brain Activity*, *supra* note 66, at 727; Daniel D. Langleben et al., *Telling Truth from Lie in Individual Subjects with Fast Event-Related fMRI*, 26 HUM. BRAIN MAPPING 262 (2005) [hereinafter Langleben et al., *Telling Truth from Lie*]; Phan et al., *supra* note 35, at 164.

⁷⁶ The fMRI playing-card paradigm was developed in an early study by Daniel Langleben’s group at the University of Pennsylvania. In general, the Guilty Knowledge Test (GKT), also called the Concealed Information Test, is one of two prevalent methods used in polygraphy (the other is the Control Question Test). In the GKT, the subject is asked questions about a fact or detail of the crime that would be known only to someone involved. The theory is that a guilty subject will show a greater physiological response to these questions than to nonrelevant questions, whereas an innocent subject will have similar reactions to both types of questions. For a description of this and other polygraph questioning techniques, see NAT’L RESEARCH COUNCIL, *supra* note 73, at 253–58; VRIJ, *supra* note 33, at 344–57 (describing and evaluating the GKT). Note that because the GKT, as used in polygraphy and in EEG “brain fingerprinting,” measures recognition, “several researchers refer to this test as a recognition test rather than a lie detection test.” VRIJ, *supra* note 33, at 344 n.1; see also Langleben et al., *True Lies*, *supra* note 72, at 356–57 (“Strictly speaking, a GKT combined with MPR [multichannel physiological recording, the measurement system used in polygraphy] is not a ‘polygraph.’”); Gamer et al., *supra* note 75, at 1288 (“[T]he GKT does not detect deception per se [T]he GKT is rather a test for recognition memory”). The GKT is often said to be more accurate than other polygraph methods, see NAT’L RESEARCH COUNCIL, *supra* note 73, at 124–25, but its potential usefulness is limited to those situations in which investigators have access to factual details that only the perpetrator would know.

playing card and a \$20 bill, and are told to memorize the card and then return it to the envelope. They are instructed that if they successfully lie and thereby fool the computer as to the identity of their playing card, they may keep the \$20 at the conclusion of the experiment.⁷⁷ Once inside the scanner, subjects are shown a series of photographs of cards, including the card in the envelope (the Lie card), and asked whether it is their card.⁷⁸ According to instructions, subjects deny having the Lie card and give truthful responses to questions about other cards shown. Each of the five studies that employed some variation of the playing-card paradigm reported statistically significant differences between the lie and truth responses, with the lie responses generally corresponding to increased activation in several brain areas thought to be involved in executive cognitive functions such as inhibition control, attention, and working memory. As summarized in one such paper, “lie

Another neuroscientific deception-detection technique that relies heavily on the guilty knowledge paradigm is the so-called “brain fingerprinting” method, which uses electrodes to measure certain changes in electrical activity in the brain known as ERPs. This technique is regarded with some skepticism in the scientific and legal communities, in large part because it has been marketed by one researcher as a foolproof, 100% accurate lie detector and because this individual claims that his algorithms are proprietary and thus his results cannot be peer-reviewed. See Mertens & Allen, *supra* note 73, at 287 (noting that “Farwell patented an ERP-based deception detection procedure and apparatus and commercially developed this deception detection procedure, which he termed ‘brain fingerprinting,’” and that he has “claimed that this approach is 100% accurate despite little empirical evidence . . . [and] despite strong contradictory evidence against one such claim of innocence” (citations omitted)). More recently, however, other researchers have begun investigating EEG and deception, and these papers have been published in peer-reviewed journals. See Ming Ann Lui et al., *supra* note 62; Mertens & Allen, *supra* note 73, at 286; Ray Johnson Jr. et al., *The Self in Conflict: The Role of Executive Processes During Truthful and Deceptive Responses About Attitudes*, 39 *NEUROIMAGE* 469 (2008); Tim R.H. Cutmore et al., *An Object Cue is More Effective Than a Word in ERP-Based Detection of Deception*, *INT’L J. PSYCHOPHYSIOLOGY* (forthcoming 2008).

As employed in the playing-card fMRI studies, however, the GKT paradigm is intended not only to measure activations triggered by recognition of a known or salient detail, but also to detect the neural substrates engaged when the subject conceals his knowledge by lying. Thus, it functions as a lie-detection test as well as a recognition-detection test, and in theory its results should generalize across situations in which the experimenter is unaware of secret details of the relevant event, in contrast to the EEG/ERP brain fingerprinting technique. I am grateful to Professor Langleben for expanding upon this point in conversation.

⁷⁷ See Langleben et al., *Brain Activity*, *supra* note 66, at 729. Test subjects are also told that if they lie about any other card (e.g., the “Nontarget,” “Control,” and “Truth” cards), they will forfeit the reward, thus ensuring a truth response to all but the “Lie” card. *Id.* Later variations of the paradigm include giving the subject the choice of which of two hidden cards will be the Lie and Truth cards, in order to moderate somewhat the experimental limitation of “lack of free choice to lie or tell the truth on specific items,” see Langleben et al., *Telling Truth from Lie*, *supra* note 75, at 270; having the subjects put the cards in their pockets, to more closely simulate a “‘real life’ experience,” see Phan et al., *supra* note 35, at 165–66; removing the monetary incentive, and modifying the experimental design so as to induce performance anxiety in subjects and thus more closely approximate and compare polygraph-type responses, see *id.* at 165.

⁷⁸ Subjects respond to these yes-or-no questions by pushing one of two buttons on a finger pad. There is also a control condition, in which they are asked, while looking at one of a number of random cards, “Is this the 10 of spades?” See, e.g., Langleben et al., *Brain Activity*, *supra* note 66, at 729.

appears to be a more working memory-intensive activity, characterized by increased activation of the [brain area] implicated in response selection, inhibition, and generation.”⁷⁹

2. Increasing Ecological Validity

The second group of studies also investigated brain response while concealing known information, but in these the known facts had more “real world” validity than in the playing-card experiments.⁸⁰ In two of these experiments,⁸¹ money was hidden under certain objects and was discovered by the subjects. The subjects were then placed inside the scanner, shown photos of the various objects, and instructed to variously give truthful or deceptive responses when asked about the location of the money.⁸² In two other studies, the subjects participated in a mock crime scenario and then answered questions about their involvement.⁸³ These studies “report predominantly right-sided frontal activation being greater during deception than truthful responding.”⁸⁴

⁷⁹ Langleben et al., *Telling Truth from Lie*, *supra* note 75, at 271; *see also* Phan et al., *supra* note 35, at 169 (“Using a modified version of the GKT, we observed specific activation of [specific areas of the prefrontal cortex] during lying, relative to both truth and control trials. These findings support the findings in prior functional anatomical studies of deception, and support the notion that discrete prefrontal cortical regions play an important role in the generation of lies, the suppression of truth, or both.”).

⁸⁰ In general, as ecological validity increases, there is greater risk that confounding variables will influence the result. S. Schwartz, *The Fallacy of the Ecological Fallacy: The Potential Misuse of a Concept and the Consequences*, 84 AM. J. PUB. HEALTH 819 (2004).

⁸¹ *See* Kozel et al., *A Pilot Study*, *supra* note 73; Frank Andrew Kozel et al., *A Replication Study of the Neural Correlates of Deception*, 118 BEHAV. NEUROSCI. 852 (2004) [hereinafter Kozel et al., *Replication Study*].

⁸² In one study, subjects were taken to one room containing six objects (for example a hat, a bowl, and a telephone), two of which concealed \$50 bills. The subjects were instructed to lie when asked whether the money was under one of these objects, and to tell the truth about the other; they were permitted to choose which object to lie about and which to tell the truth. *See* Kozel et al., *Replication Study*, *supra* note 81, at 853. In the other study, subjects were shown two rooms, each with five objects, one of which in each room concealed a \$50 bill. One room was designated the “truth room,” and the other the “deception room.” Subjects were instructed to tell the truth about the location of the money in the truth room, and to lie about the location of the money in the deception room. *See* Kozel et al., *Pilot Study*, *supra* note 73, at 297.

⁸³ *See* F. Andrew Kozel et al., *Detecting Deception Using Functional Magnetic Resonance Imaging*, 58 BIOLOGICAL PSYCHIATRY 605 (2005) [hereinafter Kozel et al., *Detecting Deception*] (subjects participated in mock crime, stealing either a ring or a watch); Mohamed et al., *supra* note 73, at 680 (subjects in the “guilty” condition participated in a mock shooting, firing a gun with blanks inside a hospital). Although this latter paper is less useful because its authors “do not formally contrast (statistically compare) deceptive and truthful responding in their report,” Spence, *Devil’s Advocate*, *supra* note 33, at 23; they do conclude that “there are specific areas of brain function that may be used to dissociate the processes of deception and truth-telling,” and thus their results appear to be broadly consistent with those of the other studies. Mohamed et al., *supra* note 73, at 686.

⁸⁴ Spence, *Devil’s Advocate*, *supra* note 33, at 23 (citing Kozel et al., *Pilot Study*, *supra* note 73; Kozel et al., *Replication Study*, *supra* note 81; Kozel et al., *Detecting Deception*, *supra* note 83).

3. *Malingering*

The third group of studies focused on a different type of deception, referred to in the psychological literature as “malingering.”⁸⁵ Though malingering is a general term for feigned impairment of one sort or another, these brain-imaging experiments sought to discover a distinctive neural signature for one specific type of malingering: feigned memory impairment. In a total of five experiments reported in three scientific papers,⁸⁶ researchers found distinct neural patterns of activation for subjects faking memory impairment as compared to simply answering incorrectly or answering randomly.⁸⁷ As with the concealed-knowledge studies discussed above, researchers here hoped to circumvent possible countermeasures to the existing non-imaging testing formats⁸⁸: “Because faking cerebral activity to avoid the detection of deception is not feasible, it was our speculation that the patterns of brain activation in

⁸⁵ “[M]alingering is the intentionally false and fraudulent simulation or exaggeration of physical or mental disease, or other defects.” Tatia M.C. Lee et al., *Lie Detection by Functional Magnetic Resonance Imaging*, 15 HUM. BRAIN MAPPING 157, 157 (2002) [hereinafter Lee et al., *Lie Detection*]. There is extensive literature on the detection of malingering using various behavioral and psychological markers. See, e.g., ASSESSMENT OF FEIGNED COGNITIVE IMPAIRMENT: A NEUROPSYCHOLOGICAL PERSPECTIVE (Kyle Brauer Boone ed., 2007); Harold V. Hall et al., *Detecting Malingering and Deception in Forensic Evaluations*, in FORENSIC PSYCHOLOGY AND NEUROPSYCHOLOGY FOR CRIMINAL AND CIVIL CASES 93–130 (Harold V. Hall ed., 2008) [hereinafter FORENSIC PSYCHOLOGY]; Harold V. Hall et al., *Deception in Criminal Contexts*, in FORENSIC PSYCHOLOGY, *supra*, at 289–328; Harold V. Hall et al., *Malingered Pain and Memory Deficits in Civil-Forensic Contexts*, in FORENSIC PSYCHOLOGY, *supra*, at 487–526; Kenneth A. Flowers et al., *Chance Guessing in a Forced-Choice Recognition Task and the Detection of Malingering*, 22 NEUROPSYCHOLOGY 273 (2008); Andrew A. Swihart et al., *Inability of the Rarely Missed Index to Identify Simulated Malingering Under More Realistic Assessment Conditions*, 30 J. CLINICAL & EXPERIMENTAL NEUROPSYCHOL. 120 (2008). See generally Edward J. Imwinkelried, *The Case Against Abandoning the Search for Substantive Accuracy*, 38 SETON HALL L. REV. 1031, 1033–42 (2008) (summarizing the various techniques currently used to detect malingering). Detection of malingering has assumed great importance because of its potential application across a variety of legal and nonlegal situations, including assessment of competency and insanity in the criminal context, pain and injury in tort cases, and accurate diagnosis and treatment in the mental health context.

⁸⁶ See Jeffrey N. Browndyke et al., *Neuroanatomical Correlates of Malingered Memory Impairment: Event-Related fMRI of Deception on a Recognition Memory Task*, 22 BRAIN INJURY 481 (2008); Lee et al., *Lie Detection*, *supra* note 85; Tatia M.C. Lee et al., *Neural Correlates of Feigned Memory Impairment*, 28 NEUROIMAGE 305 (2005) (reporting three experiments).

⁸⁷ Where individuals purposely feign memory impairment, they tend to employ certain strategies to avoid detection. As noted by Lee and colleagues, in a behavioral forced-choice test often used to detect feigned memory impairment, “performance below chance at a low probability [is] interpreted as involving the deliberate choosing of incorrect answers.” Lee et al., *Lie Detection*, *supra* note 85, at 158. Behavioral studies done in connection with fMRI research by Lee and colleagues also indicated that “the most popular cognitive strategy adopted by a majority of individuals faking memory impairment was to calculate the proportion of correct to incorrect responses by answering most of them, but not all, incorrectly.” *Id.*

⁸⁸ Though existing paradigms are considered fairly robust, “astute liars could still fake testing behavior once they understand the design of the measure.” *Id.* at 157.

malingers would provide unique markers for the detection of deception.”⁸⁹ In these studies, researchers “found widespread activation among frontal and posterior cortical regions that were greater during malingered inaccuracy than truthful, accurate, responses.”⁹⁰

4. *Autobiographical Lies*

Finally, the last group, comprised of five studies by three different research teams, compared subjects’ brain activation as they responded either truthfully or falsely to questions about autobiographical events.⁹¹ In the earliest such experiment—the first published study to use fMRI in connection with lie detection⁹²—a group of subjects first completed a questionnaire about routine activities they had done earlier that day.⁹³ They were then asked the same questions while inside the scanner.⁹⁴ Later studies extended and refined the research paradigm, for example differentiating between spontaneous versus rehearsed lies and isolated versus contextual lies,⁹⁵ comparing autobiographical and non-autobiographical content,⁹⁶ permitting subjects some choice as to whether and when to lie,⁹⁷ and examining lies about embarrassing events.⁹⁸ As with the other studies, those in this group found increased activation in portions of the prefrontal cortex for deceptive as compared to truthful responses. In addition, lies about autobiographical details elicited a greater

⁸⁹ *Id.*

⁹⁰ Spence, *Devil’s Advocate*, *supra* note 33, at 23.

⁹¹ See Ganis et al., *supra* note 73, at 830; Nuñez et al., *supra* note 64; Sean A. Spence et al., *A Cognitive Neurobiological Account of Deception: Evidence from Functional Neuroimaging*, 359 PHIL. TRANSACTIONS ROYAL SOC’Y LONDON B: BIOLOGICAL SCI. 1755, 1759–60 (2004) [hereinafter Spence et al., *A Cognitive Neurobiological Account*] (reporting preliminary results); Spence et al., *Anatomical Correlates*, *supra* note 69, at 2849; Spence et al., *Speaking of Secrets*, *supra* note 62, at 1411; see also Abe et al., *supra* note 74, at 287 (study using PET imaging to investigate activation correlated with honest and deceptive responses to autobiographical questions).

⁹² Spence et al., *Anatomical Correlates*, *supra* note 69. It should be emphasized that this first fMRI lie-detection study was published *only seven years ago*; the amount—and rate—of progress in the area is rather remarkable.

⁹³ Examples included whether they had made the bed or taken a pill. See *id.* at 2849.

⁹⁴ Subjects had been instructed to tell the truth or to lie depending on the color of the answer prompt; “[o]ver two experimental runs, each subject answered every question once with the truth and once with a lie.” *Id.* at 2849–50.

⁹⁵ See Ganis et al., *supra* note 73, at 830.

⁹⁶ See Nuñez et al., *supra* note 64.

⁹⁷ See Spence et al., *Speaking of Secrets*, *supra* note 61, at 1411, 1415.

⁹⁸ See *id.*; see also Spence et al., *A Cognitive Neurobiological Account*, *supra* note 91. Here, subjects who knew and trusted the researchers were asked to describe in writing two embarrassing events from their past, about which one might expect someone to lie if asked. These were then used to construct the questions used in the experiment. See *id.*

difference in activation than lies about non-autobiographical details,⁹⁹ and well-rehearsed and contextual lies elicited different patterns from spontaneous and non-contextual lies.¹⁰⁰

C. Recent Advances

Several recent advances in the field of brain-imaging lie detection are especially noteworthy. Perhaps most important in terms of the potential for applied use,¹⁰¹ there has been some preliminary success in distinguishing truth from lies in individual subjects.¹⁰² While most studies have thus far reported results aggregated across groups of subjects, a few recent studies have been successful at distinguishing lies from truth at the individual level.¹⁰³ In 2005, Dr. Langleben's group at the University of Pennsylvania published "the first reported formal quantification of the accuracy of fMRI-based methods for the detection of deception at the within-subject, single-event level."¹⁰⁴ Using the playing-card paradigm described above, researchers were able to achieve 78% accuracy in classifying individual events within a single subject as either truth

⁹⁹ See Nuñez et al., *supra* note 64. Because prior brain-imaging work had shown that self-relevant tasks activated brain regions associated with emotional response, this study sought to investigate "the emotive aspects of deceptive behavior." *Id.* at 268. The results were significant and distinct, and the activations interestingly tracked those seen in recent work on neural responses to personal versus impersonal moral dilemmas. *See id.* at 275, 277 (discussing J.D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCIENCE 2105 (2001)). In addition, subjects in the Nuñez study completed a personality trait questionnaire so that variations in brain activation might be associated with different personality types that arguably would "show a propensity toward deceptive behavior." *Id.* at 268. The authors concluded in this respect that "certain measurable personality traits . . . show a relationship to patterns of neural activity within" specific brain regions. *Id.* at 276.

¹⁰⁰ See Ganis et al., *supra* note 73. This study was a preliminary attempt to demonstrate that different types of lies show distinctive activation patterns. *Id.* at 835.

¹⁰¹ "The clinical value of fMRI in lie detection will be determined by the ability to detect deception in individual subjects, rather than group averages." Davatzikos et al., *supra* note 75, at 663 (abstract).

¹⁰² In addition to the fMRI studies reporting individualized results, a MEG study, which has the benefit of high temporal resolution, was also able to detect deceptive responses in individual subjects on a trial-by-trial basis. *See* Anil K. Seth et al., *Single-Trial Discrimination of Truthful from Deceptive Responses During a Game of Financial Risk Using Alpha-Band MEG Signals*, 32 NEUROIMAGE 465, 475 (2006).

¹⁰³ This is a trend not only in lie detection but in other cognitive neuroscience experiments. "For example, initial fMRI studies of motor and language areas were only able to report group analysis results. With time and refinement, fMRI is now used pre-surgically (which requires good individual predictive ability) to aid in surgical guidance." Kozel et al., *Pilot Study*, *supra* note 73, at 303; *see also* Kozel et al., *Detecting Deception*, *supra* note 83, at 605 (discussing fMRI more generally, and noting that "[w]ith improving technology and methodology . . . researchers have begun to be able to achieve reliable results at the individual level for cognitive tasks").

¹⁰⁴ Langleben et al., *Telling Truth from Lie*, *supra* note 75, at 270.

or lie.¹⁰⁵ Furthermore, because of differences noted in the accuracy of classification depending on the salience of the truth condition,¹⁰⁶ the paper suggests “that the accuracy achieved here is a low estimate of the potential accuracy of lie-detection with fMRI. . . . A task of higher ecological relevance and risk may induce larger effects and have higher accuracy.”¹⁰⁷

In a related paper using the same experimental data, Dr. Langleben’s research team was able to achieve nearly 100% accuracy in classifying truth and lie conditions for individual subjects,¹⁰⁸ and also succeeded in predicting truth or lie in previously unseen data at a rate of 88.6%.¹⁰⁹ Using a newly developed and more sophisticated classification method that looked at spatial patterns of brain activity rather than individual loci,¹¹⁰ these researchers were

¹⁰⁵ The predictive model had a sensitivity of over 76% and a specificity of over 79%. *Id.* at 267. The model was tested on four new participants to investigate whether it could be used predictively for individuals not included in the original data from which the model was generated. For these individuals, the model predicted lie versus truth with 76.5% accuracy. *Id.*

¹⁰⁶ The “truth card” in the envelope was considered more salient than the “repeat distractor” card, which in this experiment was the two of hearts. There were also a series of “variant distractor” cards, which were the remaining non-face cards of all suits. *See id.* at 263, 265.

¹⁰⁷ *Id.* at 270.

¹⁰⁸ *See* Davatzikos et al., *supra* note 75, at 668.

¹⁰⁹ *See id.* at 665–67.

¹¹⁰ *See generally* Kenneth A. Norman et al., *Beyond Mind-Reading: Multi-Voxel Pattern Analysis of fMRI Data*, 10 TRENDS COGNITIVE SCI. 424 (2006) (review article describing “a different approach to fMRI analysis, where—instead of focusing on individual voxels—researchers use powerful pattern-classification algorithms, applied to multi-voxel patterns of activity, to decode the information that is represented in that pattern of activity”). After surveying a range of applications of this innovative method, which the authors label “going beyond mind-reading,” *id.* at 427, they conclude that improvements in spatial resolution of fMRI and in computer speed, as well as more researchers working with pattern analysis, augur a “bright future” in terms of what fMRI pattern analysis will be able to do, *see id.* at 429. *See also* Shinkareva et al., *supra* note 16, at e1394 (explaining that the new approach “uses machine learning techniques to identify the neural patterns of brain activity underlying various thought processes”).

Three recently published studies illustrate the potentially revolutionary implications of computer patterning analysis. In one study, researchers trained a computer model using a “trillion word text corpus” that mapped the frequency of connections among words. Mitchell et al., *supra* note 16, at 1191. The model was then able to “predict[] fMRI neural activity well enough that it [could] successfully match words it ha[d] not yet encountered to their previously unseen fMRI images, with accuracies far above those expected by chance.” *Id.* In other words, the computer was able to predict quite accurately from a subject’s brain scan what word the person was thinking about. According to the authors, the “results establish a direct, predictive relationship between the statistics of word co-occurrence in text and the neural activation associated with thinking about word meanings.” *Id.* In another study, researchers took what they suggest is the first step toward “a general brain-reading device that could reconstruct a picture of a person’s visual experience at any moment in time.” Kay et al., *supra* note 16, at 352. Using a model generated by having subjects view 1,750 natural images while inside an fMRI scanner, the experimenters then presented 120 novel images and were able to predict quite accurately what image was being viewed based solely on the subject’s brain activation patterns. *Id.* at 353. A third study reported success in predicting which of ten objects (presented as line drawings) a subject was

able to “train” a computerized classifier both to classify individual events within the data set and to predict truth or lie in a novel subject. As they point out, this latter “approach is also more practical, because it does not require that the classifier be first trained on the person under study, before it can be applied for testing purposes.”¹¹¹ This feature is obviously crucial when imagining a method that would be useful in the typical criminal or civil litigation (or investigatory) context.

One of the mock crime scenario studies, described above, also had considerable success at differentiating truth from deception at the individual level.¹¹² In a study that has perhaps the greatest ecological validity in terms of practical lie detection in a crime or tort context,¹¹³ investigators achieved a very high accuracy rate for predicting which of two objects—a ring or a watch—had been taken, after analyzing fMRI data.¹¹⁴ The authors view this experiment as a “first step in developing this technology [which] will likely be improved with additional advances in imaging or statistical analysis.”¹¹⁵ Furthermore, they suggest that future studies that combine neuroimaging with other measures might further “improve the accuracy of detecting deception.”¹¹⁶

viewing, even where the subject was novel to the model (i.e., the computer was not trained on the particular subject). See Shinkareva et al., *supra* note 16, at 3–7.

¹¹¹ Davatzikos et al., *supra* note 75, at 667.

¹¹² See Kozel et al., *Detecting Deception*, *supra* note 83, at 605 (abstract). (The authors also claimed to be the first to investigate the neural correlates of deception in individuals rather than across groups; this paper became available at virtually the same time as the Langleben papers, Langleben et al., *Telling Truth from Lie*, *supra* note 75.)

¹¹³ As previously noted, subjects were taken to a room and instructed to steal either a ring or a watch, to put it in a locker with their other belongings, and then to lie to the investigator (while undergoing the scan) about stealing the item. To encourage them to attempt countermeasures, subjects were told they would receive an additional \$50 if the investigator operating the MRI machine was unable to detect during the test which object they had stolen. Kozel et al., *Detecting Deception*, *supra* note 83, at 606.

¹¹⁴ The analysis was done twice: once for a “model-building group” (n=30) and a second time for a “model-testing group” (n=31). Overall accuracy for the first group was 93%; for the second group, which represented data previously unseen by the classifier, the overall accuracy was 90%. See *id.* at 609–10.

¹¹⁵ *Id.* at 612. Another emerging technique that holds out promise in cognitive brain mapping is transcranial direct current stimulation (tDCS), a noninvasive procedure whereby the investigator uses an electrical current “to induce a transient interruption of normal brain activity in a relatively restricted area of the brain.” Vincent Walsh & Alan Cowey, *Transcranial Magnetic Stimulation and Cognitive Neuroscience*, 1 NATURE REV. NEUROSCI. 73, 74 (2000). By temporarily “turning off” different parts of the brain, researchers can discover which areas are implicated in different cognitive tasks, much the same as natural experiments involving individuals with injuries to specific parts of the brain. See *id.* The technique has recently been employed to investigate deception. See Priori et al., *supra* note 46, at 451–52 (experimentally influencing the production of different types of lie responses, thereby suggesting that these have distinct neural mechanisms).

¹¹⁶ Kozel et al., *Detecting Deception*, *supra* note 83, at 612.

Other studies, too, have combined behavioral or peripheral nervous measures with fMRI to further refine deception theory and data.¹¹⁷

In addition to these technological and statistical advances, the study paradigms have been steadily improving as cognitive neuroscientists integrate new data into existing models of deception. As one group points out:

One of the strengths of contemporary theory in cognitive neuroscience is that it distinguishes subtypes of a given function. For example, “memory” may be decomposed into working, episodic, semantic memory, and so on. . . . [W]e applied the same approach to study deception and asked what types of processing differences might distinguish different types of lies.¹¹⁸

Following this insight, the more recent studies have begun to refine the experimental paradigms and to distinguish among subtypes of deception. As previously noted, the work of the various research teams had already shown a tendency to focus on distinct subtypes of deception, with some studies looking at feigned memory impairment, some at deception regarding autobiographical data, some at deception regarding past actions, and the like. As data accumulates, researchers have begun to develop their hypotheses and study designs based upon increasingly sophisticated theories of deception. In addition, they have been able to modify their experiments to isolate genuinely deceptive responses from possibly confounding factors.¹¹⁹

In one recent study, subjects were permitted to choose when to tell the truth and when to lie so as to address the potentially problematic effect of “experimenter-sanctioned cued-deception”—subjects being told by the experimenter when to lie.¹²⁰ In another, researchers sought to simulate the stress caused by fear of detection typically associated with polygraphy by using a novel real-time scanning technology and informing subjects that the experimenter would see their results instantaneously.¹²¹ They surmised that

¹¹⁷ See Gamer et al., *supra* note 75 (combining fMRI, reaction time (RT), and skin conductance measurements); Kozel et al., *Pilot Study*, *supra* note 73 (combining BOLD fMRI and electrodermal activation measures). Kozel et al. have suggested that “combining fMRI . . . with the polygraph offers the potential for increasing the sensitivity and specificity of a combined fMRI/polygraph machine” that might reliably detect deception on an individual level. *Id.* at 303.

¹¹⁸ Ganis et al., *supra* note 73, at 830–31.

¹¹⁹ For example, one recent study used voice responses rather than button pushes to eliminate possible noise caused by neural correlates of motor responses. See Spence et al., *Speaking of Secrets*, *supra* note 62, at 1412–13.

¹²⁰ See *id.*

¹²¹ See Phan et al., *supra* note 35.

these subjects might experience an unconscious, somatic response associated with knowledge that their conduct was wrong and that it was being socially observed, which might be absent where results were to be analyzed at a later time.¹²² Another experiment contrasted spontaneous and memorized lies, as well as isolated and contextualized lies.¹²³ The researchers hypothesized that these various species of deception would activate different brain areas, with spontaneous lies expected to rely on episodic memory, semantic memory, or both; memorized lies expected to rely on retrieval from episodic memory; and contextualized lies increasing working memory activation.¹²⁴ These and other experiments continue to advance deception theory and detection.

Accurate detection of when a subject is being deceptive is not equivalent to “mind reading,” and the above-described experiments do not come close to making the black box of the mind an open book. But they do open a small window into the inner workings of the brain, which in turn could open a window into the black box of the jury room. This possibility is addressed in the next Part.

II. INSIDE THE SECOND BLACK BOX: WATCHING THE JURY DECIDE

The balance of this Article assumes that the justice system might realistically be presented with an extremely reliable and practical lie detection method based on brain-imaging technology. Under those circumstances, it would be necessary to ask whether such a technique should be admissible, even given a very high degree of accuracy.¹²⁵ There is reason to suspect that the system would remain resistant to a very good—indeed, even a foolproof—lie detector. This Part looks more closely at the black box of the jury room to

¹²² *Id.* at 170.

¹²³ *See* Ganis et al., *supra* note 73.

¹²⁴ *Id.* at 831.

¹²⁵ As one scholar has asserted,

Various devices have been proposed over the years to determine the truthfulness of a subject's account, ranging from “truth serum” to polygraphs that measure blood pressure and pulse to counting the rate of eye blinks. *But even if we were satisfied with the scientific reliability of a particular method, it is inconceivable that we would permit its use at trial to test the truthfulness of a witness' testimony.*

Mark S. Brodin, *Behavioral Science Evidence in the Age of Daubert: Reflections of a Skeptic*, 73 U. CIN. L. REV. 867, 925 (2005) (citations omitted) (emphasis added). Professor Michael Pardo has analyzed the Fourth and Fifth Amendment issues that would be presented by a reliable neuroimaging-based lie detector. *See* Pardo, *supra* note 27. In contrast, this Article focuses on the potential implications for the jury system of a reliable neuroimaging-based lie detector.

illuminate the multiple functions contained inside that are hidden from public scrutiny by a host of procedural and evidentiary rules. Primary among these is the principle that determining witness credibility is the sole province of the jury, and that evidence that threatens to invade this sacred realm is excludable for that reason alone; admission of reliable lie detection evidence would directly challenge this entrenched principle. Perhaps even more important, such evidence would indirectly undermine a second bedrock principle: the process of jury decision making must remain cloaked in complete secrecy. This Part first examines the jury's traditional and jealously guarded role as the exclusive fact finder and sole arbiter of credibility. It then suggests various scientific and social forces that, in the face of a reliable lie detection technology, are likely to increase pressure on the system and, perhaps, to crack open the tight lid on the admissibility of reliable credibility evidence. Finally, this Part addresses the broader implications for the jury system of a machine capable of telling truth from lies.

A. *Shutting the Lid: Protecting the Jury as Sole Fact Finder*

Though subject to various exceptions and complexities,¹²⁶ it is commonplace that the “controlling distinction between the power of the court and that of the jury is that the former is the power to determine the law and the latter to determine the facts.”¹²⁷

¹²⁶ See, e.g., James B. Thayer, “*Law and Fact*” in *Jury Trials*, 4 HARV. L. REV. 147 (1890) (arguing that what we call “fact” is that which we believe juries should decide, and what we call “law” is that which we believe judges should decide, but that the line between the categories is otherwise fuzzy); see also HENRY M. HART, JR. & ALBERT M. SACKS, *THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF LAW* 344–60 (William N. Eskridge & Philip P. Frickey eds., 1994) (proposing that the issue be discussed with reference to judge questions and jury questions rather than law and fact).

¹²⁷ *Dimick v. Schiedt*, 293 U.S. 474, 486 (1935); see, e.g., Risinger, *supra* note 38, at 1290 (“In the standard version of the rationalist model [of the trial], juries decide facts.”); Simmons, *supra* note 27, at 1013 (“In our judicial system, the jury functions as the ultimate and exclusive finder of fact.”). Federal Rule of Evidence 104 governs the allocation of authority between judge and jury in the federal courts. The rule provides:

Preliminary questions concerning the qualification of a person to be a witness, the existence of a privilege, or the admissibility of evidence shall be determined by the court [However, when the relevancy of evidence depends upon the fulfillment of a condition of fact, the court shall admit it upon, or subject to, the introduction of evidence sufficient to support a finding of the fulfillment of the condition.]

FED. R. EVID. 104(a)–(b). The reservation to the jury of questions of conditional relevance was thought necessary to avoid impinging upon the jury's protected authority to decide ultimate factual issues at trial. See FED. R. EVID. 104 advisory committee's note. As to the power of juries to decide the law, see *infra* Part III.

The American system appears to go to great lengths to guard the jury's fact-finding function.¹²⁸ Courts wax poetic about this jury role;¹²⁹ a recent spate of Supreme Court opinions has reemphasized that the jury's fact-finding function in the criminal context has significant constitutional moorings: the Sixth Amendment right to a trial by jury *is* the right to have the jury, rather than the judge, determine certain facts.¹³⁰ Furthermore, evaluating the

¹²⁸ See, e.g., *Chauffers, Teamsters & Helpers Local No. 391 v. Terry*, 494 U.S. 558, 565 (1990) (quoting *Beacon Theatres, Inc. v. Westover*, 359 U.S. 500, 501 (1959)); *Dimick*, 293 U.S. at 486 (discussing Seventh Amendment right to jury trial in civil cases, and stating that “[m]aintenance of the jury as a fact-finding body is of such importance and occupies so firm a place in our history and jurisprudence that any seeming curtailment of the right to jury trial should be scrutinized with the utmost care”); *United States v. Adams*, 271 F.3d 1236, 1246 (10th Cir. 2001) (discussing “the jury’s vital and exclusive function to make credibility determinations”). It may be that some of this protection of the jury’s exclusive role is more apparent than real. Certainly in the civil context the jury’s power to determine facts has been steadily eroded. See Suja A. Thomas, *The Seventh Amendment, Modern Procedure, and the English Common Law*, 82 WASH. U. L.Q. 687 (2004) [hereinafter Thomas, *The Seventh Amendment*] (detailing rise of procedures such as summary judgment, special verdicts, and remittitur that have weakened the fact-finding role of the civil jury). In the criminal context, too, courts have assumed increased fact-finding authority though the jury’s authority remains stronger in the criminal context. See Suja A. Thomas, *Judicial Modesty and the Jury*, 76 U. COLO. L. REV. 767 (2005) [hereinafter Thomas, *Judicial Modesty*] (arguing that courts have exhibited “judicial modesty” in policing the bounds of judicial power in relation to that of the jury in the criminal context, whereas they have not exercised such restraint in the civil context, but rather have allowed judicial power to increase at the expense of the jury). Thayer noted more than 100 years ago, in discussing the jury’s authority to decide questions of fact, that

in theory, the judges have almost always paid homage to the jury’s separate and independent right. Seldom have they failed to do that. Yet the judges supervise and moderate their action, and herein lies one of the most searching and far-reaching grounds of judicial control,—that of keeping the jury within the bounds of reason.

Thayer, *supra* note 126, at 167.

¹²⁹ The Supreme Court has repeatedly stated,

The right of trial by jury is of ancient origin, characterized by Blackstone as “the glory of the English law” and “the most transcendent privilege which any subject can enjoy”; and, as Justice Story said, “. . . the Constitution would have been justly obnoxious to the most conclusive objection if it had not recognized and confirmed it in the most solemn terms.” With, perhaps, some exceptions, trial by jury has always been, and still is, generally regarded as the normal and preferable mode of disposing of issues of fact in civil cases at law as well as in criminal cases.

Dimick, 293 U.S. at 485–86 (internal citations omitted); see also *United States v. Stromberg*, 179 F. Supp. 278, 280 (S.D.N.Y. 1959) (“The most important function served by a jury is in bringing its accumulated experience to bear upon witnesses testifying before it, in order to distinguish truth from falsity. Such a process is of enormous complexity, and involves an almost infinite number of variable factors. It is the basic premise of the jury system that twelve men and women can harmonize those variables and decide, with the aid of examination and cross-examination, the truthfulness of a witness.”).

¹³⁰ See *Crawford v. Washington*, 541 U.S. 36, 68 (2004) (holding that, under Sixth Amendment Confrontation Clause, testimonial statements by witnesses offered against the accused must be presented to the jury in the form of live testimony rather than as hearsay); *Apprendi v. New Jersey*, 530 U.S. 466, 490 (2000)

credibility of witnesses lies at the heart of this jury function. “It is axiomatic that assessing the credibility of witnesses is the sole prerogative of the jury. Indeed, it has been said that the genius of the jury trial system is to have twelve laypersons perform this task”¹³¹ A range of procedural and evidentiary rules are typically justified as necessary to guard the province of the jury to find facts and, especially, to evaluate credibility. At trial, evidence that tends to impinge upon this singular jury function has usually been viewed skeptically by courts and is frequently excluded. On appeal, reviewing courts will uphold a jury verdict if the jury *might* have found facts that support the verdict, even if that possibility seems remote.¹³² Finally, rules that protect the secrecy of jury deliberations and the finality of verdicts work in tandem with these credibility doctrines to preserve the possibility that a jury verdict was based on a correct application of the law to the facts as found by the jury.¹³³

(holding that, pursuant to the Sixth Amendment guarantee of trial by jury, “any fact that increases the penalty for a crime beyond the statutory maximum must be submitted to a jury, and proved beyond a reasonable doubt”); *see also* *Cunningham v. California*, 127 S. Ct. 856 (2007) (reaffirming constitutional right to jury decision of facts necessary to conviction or sentencing); *Blakely v. Washington*, 542 U.S. 296, 303–05 (2004) (holding that Washington sentencing statute violated defendant’s Sixth Amendment right to trial by jury where judge subjected defendant to “exceptional” sentence based on judicial finding of deliberate cruelty based on facts not found by a jury nor admitted by defendant); *United States v. Booker*, 543 U.S. 220, 234–37 (2005) (invalidating federal sentencing guidelines on same grounds); *United States v. Gaudin*, 515 U.S. 506, 511 (1995) (holding that a criminal defendant has a constitutional “right to demand that a jury find him guilty of all the elements of the crime for which he is charged”); *Seaman*, *supra* note 44, at 851–60 (discussing the *Crawford* and *Apprendi* line of cases). The right to trial by jury in criminal cases is guaranteed by the Sixth Amendment, which has been incorporated as against the States, and also under Article III of the Constitution. In the civil context too, of course, the right to a jury trial in many cases is constitutionally guaranteed under the Seventh Amendment and that jury right explicitly protects the right to jury determinations of fact. The Seventh Amendment provides: “In suits at common law . . . the right of trial by jury shall be preserved, and no fact tried by a jury, shall be otherwise re-examined in any Court of the United States, than according to the rules of the common law.” U.S. CONST. amend. VII.

¹³¹ Brodin, *supra* note 125, at 916; *see* Chet K.W. Pager, *Blind Justice, Colored Truths and the Veil of Ignorance*, 41 WILLAMETTE L. REV. 373, 375 (2005) (noting that the “axiom” that “[d]etermination of witness credibility is the exclusive domain of the jury . . . is deeply embedded in both our history and jurisprudence”).

¹³² Professor Risinger summarizes the state of the doctrine with respect to appellate review of jury fact-finding as follows:

In virtually every American jurisdiction, when the sufficiency of evidence to support a verdict is attacked, . . . [t]he party prevailing below is entitled to every inference that a reasonable jury might have made given the evidence on the record considered in its most favorable light. This essentially means accepting at face value all testimonial evidence in favor of the verdict and assuming all testimonial evidence to the contrary to have been rejected on credibility grounds.

Risinger, *supra* note 38, at 1313–14.

¹³³ *See* FED. R. EVID. 606(b) (barring jurors from testifying as to their internal deliberations or thought processes in order to impeach their verdict); *Tanner v. United States*, 483 U.S. 107 (1987) (interpreting scope of Rule 606(b) very broadly to exclude post-verdict juror testimony regarding use by jurors during trial of alcohol and illegal drugs).

Several scholars have observed that the judicial reaction against credibility expertise, and lie detection evidence in particular,¹³⁴ has been wholly out of proportion to its purported lack of scientific reliability.¹³⁵ Professor James McCall, who has analyzed the admissibility of polygraph evidence,¹³⁶ called the judicial response “resistance on an historic scale” and noted “the surprising extent and vigor of the ‘search for plausible legal objections’ to avoid consideration of the claim” that such evidence had reached a level of scientific reliability sufficient to satisfy *Daubert* and Rule 702.¹³⁷ In the wake of the Supreme Court’s decision in *Daubert v. Merrell Dow Pharmaceuticals*,¹³⁸ which rejected the *Frye* “general acceptance” test¹³⁹ under the *Federal Rules of Evidence*, several scholars predicted that courts would reevaluate their longstanding per se rejection of polygraph evidence.¹⁴⁰ Yet after some initial

¹³⁴ Other forms of expertise that implicate witness credibility include expert testimony about the reliability of eyewitness identification and psychological testimony concerning the behavior of child abuse and rape victims.

¹³⁵ Though the reliability and validity of polygraphy are far from settled, it seems clear enough that the technique is at least as reliable as several other forensic techniques that are routinely admitted. See Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. CRIM. L. & CRIMINOLOGY 985 (2005) (questioning validity and accuracy of fingerprint identification); Edward J. Imwinkelried & William A. Tobin, *Comparative Bullet Lead Analysis (CBLA) Evidence: Valid Inference or Ipse Dixit?*, 28 OKLA. CITY U. L. REV. 43 (2003) (critiquing bullet lead evidence); Michael J. Saks, *Banishing Ipse Dixit: The Impact of Kumho Tire on Forensic Identification Science*, 57 WASH. & LEE L. REV. 879 (2000); Adina Schwartz, *A Systemic Challenge to the Reliability and Admissibility of Firearms and Toolmark Identification*, 6 COLUM. SCI. & TECH. L. REV. 1 (2005) (questioning validity of toolmark identification); C.A. Stafford Smith & P.D. Goodman, *Forensic Hair Comparison Analysis: Nineteenth Century Science or Twentieth Century Snake Oil?*, 27 COLUM. HUM. RTS. L. REV. 227 (1996); William C. Thompson & Michele Nethercott, *Forensics: Lessons from Baltimore’s GSR Debacle*, CHAMPION, June 2005, at 50 (exposing problems with gunshot residue evidence).

¹³⁶ See, e.g., James R. McCall, *Misconceptions and Reevaluation—Polygraph Admissibility after Rock and Daubert*, 1996 U. ILL. L. REV. 363 [hereinafter McCall, *Misconceptions*]; Imwinkelried & McCall, *supra* note 37.

¹³⁷ James R. McCall, *The Personhood Argument Against Polygraph Evidence, or “Even If the Polygraph Really Works, Will Courts Admit the Results?”*, 49 HASTINGS L.J. 925, 925–26 (1998); see Simmons, *supra* note 27, at 1043 (“[C]ourts lag dramatically behind the rest of the country in accepting this new technology [of polygraphy], and the courts’ reluctance is based on something other than legitimate reliability concerns.”).

¹³⁸ 509 U.S. 579 (1993).

¹³⁹ See *Frye v. United States*, 293 F. 1013, 1013–14 (D.C. Cir. 1923) (upholding trial court’s exclusion of expert testimony regarding the results of a systolic blood pressure lie detection test taken by defendant, and stating that evidence based on a scientific principle or discovery should be admissible when “the thing from which the deduction is made [is] sufficiently established to have gained general acceptance in the particular field in which it belongs”).

¹⁴⁰ See, e.g., Sheila K. Hyatt, *Developments in the Law of Scientific Evidence: The Admissibility of Polygraph Evidence*, 18 J. NAT’L ASS’N ADMIN. L. JUDGES 171, 194 (1998) (“[P]olygraph evidence is likely to become more common in the courtroom”); Imwinkelried & McCall, *supra* note 37, at 1080 (“The battle over the admissibility of polygraph evidence should be waged over the issue of its scientific validity under

signs of a trend in that direction,¹⁴¹ most courts have continued to exclude polygraph evidence.¹⁴² In doing so, they frequently cite the danger that such evidence will usurp the province of the jury to determine witness credibility.¹⁴³

As Professor Ric Simmons has demonstrated, this “province of the jury” doctrine is derived from two distinct historical forebears: the prohibition against opinion testimony by witnesses, and the traditional separation and balance of powers between court and jury.¹⁴⁴ With respect to the former, witnesses were permitted to testify if they had some special expertise such that their opinions would assist the jury. However,

Daubert.”); McCall, *Misconceptions*, *supra* note 136, at 422 (“What can be predicted, however, is that polygraph evidence will now be admitted in American trial courts on a greater scale than was conceivable before *Daubert.*”); Richard H. Underwood, *Truth Verifiers: From the Hot Iron to the Lie Detector*, 84 KY. L.J. 597, 632 (1996) (“*Daubert* . . . opens the door wide to reconsideration of polygraph evidence.”).

¹⁴¹ See *United States v. Call*, 129 F.3d 1402 (10th Cir. 1997) (rejecting per se ban on polygraph in light of *Daubert*); *United States v. Cordoba*, 104 F.3d 225 (9th Cir. 1997) (same); *United States v. Posado*, 57 F.3d 428 (5th Cir. 1995) (same).

¹⁴² See, e.g., *United States v. Gill*, 513 F.3d 836, 846 (8th Cir. 2008) (“Our cases make clear polygraph evidence is disfavored.”); *United States v. Gardiner*, 463 F.3d 445, 469 n.8 (6th Cir. 2006) (“Admission of polygraph evidence is disfavored in this Circuit”); *United States v. Prince-Oyibo*, 320 F.3d 494, 501 (4th Cir. 2003) (“[T]o the extent that *Daubert*’s alteration of the legal landscape threw into doubt the viability of our per se rule against polygraph evidence, [subsequent Fourth Circuit cases] effectively resolved those doubts in favor of the rule.”); *Bloom v. People*, 185 P.3d 797, 801 (Colo. 2008) (“[The] per se ban is an evidentiary rule rooted in the concern that polygraph evidence will prejudice the jury[. . . .]”); *Thornton v. State*, 620 S.E.2d 356, 360 (Ga. 2005) (“The results of a polygraph are inadmissible except by stipulation of the parties”); *Wilkins v. State*, 190 P.3d 957, 970 (Kan. 2008) (“[R]eference to [a polygraph] examination . . . is prohibited.”); *State v. Jones*, 753 N.W.2d 677, 690 (Minn. 2008) (“Polygraph examinations are inadmissible.”); *State ex rel. Kemper v. Vincent*, 191 S.W.3d 45, 49 (Mo. 2006) (“The results of a polygraph examination generally are inadmissible in Missouri criminal trials.”); *State v. Hameline*, 188 P.3d 1052, 1055 (Mont. 2008) (“[P]olygraph results are inadmissible”); *Darling v. State*, 262 S.W.3d 913, 920 (Tex. Ct. App. 2008) (“The results of a polygraph examination are generally inadmissible for any reason because such testing is inherently unreliable.”); *Fowlkes v. Commonwealth*, 663 S.E.2d 98, 101–02 (Va. Ct. App. 2008) (“[A] polygraph examination has no proper evidentiary use” (quoting *Gray v. Graham*, 341 S.E.2d 153, 157 (Va. 1986))). Though other types of expert testimony that implicate witness credibility have gained some measure of acceptance in the courts, polygraph evidence “is the current (and perhaps final) battleground for the admissibility of evidence that was once thought to invade the province of the jury.” Simmons, *supra* note 27, at 1016.

¹⁴³ See, e.g., *United States v. Wilson*, 183 Fed. App’x 814 (10th Cir. 2006) (stating that “the credibility of witnesses is generally not an appropriate subject for expert testimony,” including testimony regarding polygraph data, because such testimony “usurps a critical function of the jury . . . , which is capable of making its own determination regarding credibility”); *United States v. Swayze*, 378 F.3d 834, 937 (8th Cir. 2004) (“When two witnesses contradict each other, juries, not polygraph tests, determine who is testifying truthfully.”); *United States v. Call*, 129 F.3d 1402, 1406 (10th Cir. 1997) (expert testimony in regard to a polygraph test “is often excluded because it usurps a critical function of the jury”); *State v. Engelhardt*, 119 P.3d 1148, 1166 (Kan. 2005) (“This rule is attributable . . . in part to protection of the jury’s role as the factfinder.”).

¹⁴⁴ Simmons, *supra* note 27, at 1017–20.

if the witness was merely drawing conclusions from the facts based on his own common sense, rather than some special skill, he was considered to be doing the jurors' job for them. . . . Thus the claim that an expert was invading the province of the jury was merely a shorthand method of arguing that the expert testimony was not based on the witness's expertise and therefore was not helpful to the jury.¹⁴⁵

Quite distinct from the witness opinion rule, courts have long used language of usurpation and invasion of the jury's province in the context of the division of authority between judge and jury. As between these two institutional actors, the jury is vested with the authority to determine adjudicative facts, including the credibility of witnesses. Both trial judges and appellate judges are expected to defer to the jury's reasonable factual determinations and to refrain from substituting their judgments on factual matters for those of the jury. Professor Simmons argues that the common usage of the "province of the jury" language in these two distinct contexts has led courts—inappropriately in his view—to conflate the rationales for the rules and to conclude that an expert witness may not opine on the credibility of another witness regardless of the reliability of her method and regardless of whether the opinion would actually be helpful to the jury.¹⁴⁶ "Thus, the concept that credibility is a topic uniquely inappropriate for expert testimony—a concept incubated for decades by the 'beyond the scope' restriction—now stands alone to justify precluding expert testimony on credibility."¹⁴⁷ As Supreme Court Justice Clarence Thomas emphatically stated in a decision that upheld a per se exclusion of polygraph evidence in military courts-martial, "A fundamental premise of our criminal trial system is that 'the jury is the lie detector.'"¹⁴⁸

¹⁴⁵ *Id.* at 1018.

¹⁴⁶ According to Professor Simmons,

This logic is problematic, because allowing an expert to testify regarding her opinion about another witness's credibility is not analogous to allowing a trial judge to state or imply an opinion on the matter (or to an appellate court substituting its own opinion on witness credibility for the trial jury's opinion). Rather, the expert witness is merely stating an opinion that, like all expert opinions and indeed all other testimony, the jury can accept, reject, or give as much weight as it sees fit.

Id. at 1021.

¹⁴⁷ *Id.* at 1029.

¹⁴⁸ *United States v. Scheffer*, 523 U.S. 303, 312 (1998) (quoting *United States v. Barnard*, 490 F.2d 907, 912 (9th Cir. 1973)) (emphasis added by *Scheffer* Court). The portion of the *Scheffer* opinion that contained the quoted language, which provided, as an alternate reason for its holding, that the per se exclusion was justified by the legitimate governmental interest in "[p]reserving the [fact finder's] core function of making

Such restrictions on the admissibility of expert testimony regarding witness credibility help to keep the black box of the jury room shut by preserving ambiguity in the evidence.¹⁴⁹ If the jury is the sole determiner of credibility, then so long as the evidence meets a bare sufficiency requirement any verdict can potentially be justified as a credibility determination by the jury.¹⁵⁰ Whether the jury *actually* decided the case on that basis is a distinct question, but the answer is obscured by the rules regarding jury secrecy. Thus, whereas the credibility restrictions work indirectly to preserve the black box quality of jury decision making, the rules regarding jury secrecy directly lock the black box.

Rules that preserve the secrecy of jury deliberations attempt to place a shroud around the jury box and prevent any information from leaking out. Rule 606(b) of the *Federal Rules of Evidence* prohibits the use of juror testimony to impeach a verdict where such testimony concerns “any matter or statement occurring during the course of the jury’s deliberations” or “the effect of anything upon that or any other juror’s mind or emotions as influencing the juror to assent to or dissent from the verdict . . . or concerning the juror’s mental processes in connection therewith.”¹⁵¹

credibility determinations in criminal trials,” was joined by only four Justices. *Id.* at 312–13. Justice Kennedy concurred in part and concurred in the judgment, but published a separate opinion in which he stated: “[I]t seems the principal opinion overreaches when it rests its holding on the additional ground that the jury’s role in making credibility determinations is diminished when it hears polygraph evidence.” *Id.* at 318 (Kennedy, J., concurring). Justice Kennedy’s concurrence was joined by Justices O’Connor, Ginsburg, and Breyer. Justice Stevens dissented; he rejected the plurality’s argument that admission of polygraph evidence would usurp the role of the jury in determining credibility. *See id.* at 336 (Stevens, J., dissenting). Several commentators have pointed out that *Scheffer* therefore reveals five Justices in support of a reevaluation of the per se rule of polygraph inadmissibility after *Daubert*. It also reveals five Justices who reject the premise that the role of the jury in assessing witness credibility requires exclusion of polygraph evidence. For an excellent discussion of the precedential value of such split Supreme Court decisions, see Michael L. Eber, Comment, *When the Dissent Creates the Law: Cross-Cutting Majorities and the Prediction Model of Precedent*, 58 EMORY L.J. 207 (2008).

¹⁴⁹ Of course, so long as the expert’s opinion could reasonably be rejected by the jury, ambiguity (though perhaps lessened) would still be preserved. A cynic might therefore suggest that some of the resistance to expert credibility evidence stems from the fact that it has most often been offered by the defendant rather than the prosecution. Indeed, at the same time that the government argues that polygraph evidence is unreliable, it continues to use it extensively in investigatory and employment contexts, among others. *See* ALFRED CUMMING, CONG. RES. SERV., POLYGRAPH USE BY THE DEPARTMENT OF ENERGY: ISSUES FOR CONGRESS 1–5 (2008), available at <http://www.fas.org/sgp/crs/intel/RL31988.pdf>.

¹⁵⁰ *Cf.* Risinger, *supra* note 38, at 1282 (arguing that each of the two dominant models of the trial “allows judges—especially appellate judges—to avoid responsibility for conviction of the factually innocent”).

¹⁵¹ FED. R. EVID. 606(b).

The rule provides exceptions, rooted in the common law, for juror testimony concerning “extraneous prejudicial information” or “outside influence.”¹⁵² As interpreted by the federal courts and many state courts,¹⁵³ Rule 606(b) thus codifies a distinction between “internal” and “external” influences on the verdict, with jurors permitted to testify post-verdict to the latter but not the former.¹⁵⁴ Within the metaphor of the black box, jurors may testify as to whether information might have entered the box having been neither approved by the court nor tested during the adversarial process.¹⁵⁵ However, as to the mysterious process that occurs once all of the data has been entered, such may not be disclosed.¹⁵⁶

Courts and scholars note several reasons that have been advanced in support of the rule preventing jurors from post-verdict impeachment; among these is the importance of secrecy in jury deliberations.¹⁵⁷ As a distinct justification for the rule,¹⁵⁸ protecting the secrecy of the jury room is said to allow “frankness and freedom of discussion and conference” among jurors,¹⁵⁹

¹⁵² *Id.* (“But a juror may testify about (1) whether extraneous prejudicial information was improperly brought to the jury’s attention, (2) whether any outside influence was improperly brought to bear upon any juror . . .”).

¹⁵³ For an appendix detailing the rules of each state, see Benjamin T. Huebner, Note, *Beyond Tanner: An Alternative Framework for Postverdict Juror Testimony*, 81 N.Y.U.L. REV. 1469, 1500–04 (2006).

¹⁵⁴ See *Tanner v. United States*, 483 U.S. 107, 117 (1987); see also Huebner, *supra* note 153, at 1480–81.

¹⁵⁵ Classic examples include jurors reading newspaper articles about the case or bringing such material into the jury room, a bailiff making a comment about the defendant or the evidence to the jurors, or a juror being subjected to a bribe or other form of tampering. See generally 27 CHARLES ALAN WRIGHT & VICTOR JAMES GOLD, FEDERAL PRACTICE AND PROCEDURE § 6075 (1990); see also, e.g., *Remmer v. United States*, 347 U.S. 227 (1954) (attempt to bribe jury foreman); *McGuire v. Howard*, 128 S.E.2d 281 (Va. 1962) (unauthorized juror visit to accident scene); *Quintana v. State*, 405 P.2d 740 (Colo. 1965) (newspaper article containing information not admitted into evidence).

¹⁵⁶ As others have noted, however, there is no general rule that precludes jurors from discussing their deliberations; Rule 606(b) prohibits only juror testimony in a proceeding challenging the legitimacy of the verdict. It is increasingly common for jurors to speak to the media or otherwise to share their accounts of deliberations in high-profile cases. See, e.g., Nicole B. Cásarez, *Examining the Evidence: Post-Verdict Interviews and the Jury System*, 25 HASTINGS COMM. & ENT. L.J. 499 (2003); Copernicus T. Gaza, Note, *Getting Inside the Jury’s Head: Media Access to Jurors After the Trial*, 12 N.Y.L. SCH. J. HUM. RTS. 311 (1995).

¹⁵⁷ The other commonly cited reasons include preventing harassment of jurors by the parties, preserving the finality of verdicts, and preventing juror fraud. See *Tanner*, 483 U.S. at 107; *Shillcutt v. Gagnon*, 827 F.2d 1155, 1158 (7th Cir. 1987); Victor Gold, *Juror Competency to Testify that a Verdict Was the Product of Racial Bias*, 9 ST. JOHN’S J. LEGAL COMMENT. 125 (1993); *Developments in the Law—Race and the Criminal Process*, 101 HARV. L. REV. 1472, 1599 (1988).

¹⁵⁸ The various justifications may overlap, as when secrecy helps to preserve verdict finality. In that case, secrecy is not a distinct justification for the rule, but simply a consequence of the rule, which helps to further the value of finality.

¹⁵⁹ *McDonald v. Pless*, 238 U.S. 264, 268 (1915).

which in turn permits the jury system to function effectively.¹⁶⁰ In making this argument in favor of Rule 606(b), its supporters suggest that the very system might be threatened in the absence of a secrecy shield: in effect, the argument is that revealing the workings of the black box could destroy the system of jury decision making entirely.¹⁶¹

This argument, in turn, has a number of distinct components. The first relates to the importance of full and frank deliberation. According to this line of reasoning, the secrecy protection operates in the nature of a privilege, and promotes communication among jurors in the same way that other evidentiary privileges are believed to promote candor and openness.¹⁶² If opening the process of jury deliberation to scrutiny were to have the effect of diminishing or altering the quality of the deliberation, then it could fairly be said that the black box aspect of the jury is necessary to the integrity of the process itself and, in turn, justifies a secrecy rule for its own sake.

A separate thread that runs through some of the secrecy discussions is epitomized by the Supreme Court's suggestion in *Tanner v. United States* that "[i]t is not at all clear . . . that the jury system could survive such efforts to perfect it" as would be required by allowing inquiries into the internal processes of juror deliberations.¹⁶³ The implication of this language is that knowing too much about how a jury verdict is made would, as with sausages or laws, cause the public to reject the product.¹⁶⁴ According to this argument, the legitimacy of the jury system would be threatened were the black box to be opened to public view.

The manner in which the black box of the jury room would be cracked open by the diminishment of the jury's fact-finding or credibility-determining role, however, is different in kind from that raised by Rule 606(b) inquiries. Unlike revelation of the content of deliberations or the substantive reasons for jury decisions, the view inside the jury box that might be occasioned by a

¹⁶⁰ *Tanner*, 483 U.S. at 124–25.

¹⁶¹ In a much-quoted passage, Justice O'Connor wrote: "There is little doubt that postverdict investigation into juror misconduct would in some instances lead to the invalidation of verdicts reached after irresponsible or improper jury behavior. It is not at all clear, however, that the jury system could survive such efforts to perfect it." *Id.* at 120.

¹⁶² This was Dean Wigmore's approach. See Gold, *supra* note 157, at 135 (critiquing the privilege rationale as both over- and under-inclusive).

¹⁶³ *Tanner*, 483 U.S. at 120.

¹⁶⁴ See Clifford Holt Ruprecht, Comment, *Are Verdicts, Too, Like Sausages?: Lifting the Cloak of Jury Secrecy*, 146 U. PA. L. REV. 217, 217 (1997) (quoting a line attributed to Otto von Bismark: "If you like laws and sausages, you should never watch either one being made").

foolproof lie detector would be one of function rather than content. To the extent that a jury's view of its own role would be shaped by its assessment of the level of certainty of the evidence before it, jurors might have some concern about public disagreement with a nullificatory verdict. However, it is difficult to see why such concerns would be any greater than, or materially different from, concerns that the public might disagree with its decision on the facts.

With respect to the legitimacy concern, the question is whether the public would be less likely to accept jury verdicts were the legislative role of the jury to become increasingly apparent. Arguably, public acceptability of verdicts should correspond with the degree of fit between what we say the jury should do and what it actually appears to do. If that is the case, then an insistence that the jury acts solely as a fact finder in the face of a decreasing fact-finding role would diminish, rather than enhance, its institutional legitimacy. In any particular case, and given reduced room for doubt about historical facts, an instance of nullification would be likely to increase legitimacy concerns where the system insisted on characterizing nullification as lawlessness.¹⁶⁵

B. Pressure Building Outside the Box

As described above, there are thus some very powerful forces tending to resist evidence that would expose the jury decision-making process to close scrutiny. At the same time, there are contrary forces operating in the other direction.¹⁶⁶ This section considers the potential impact on the criminal justice system of a confluence of recent developments that, according to this Article's hypothesis, may soon include the development of highly accurate brain-imaging lie detection techniques. These recent phenomena include powerful and highly accurate scientific forensic techniques such as DNA typing,¹⁶⁷ along with the numerous and widely publicized instances of exoneration of factually innocent persons who had falsely been convicted, and research into

¹⁶⁵ Cf. Sherman J. Clark, *The Courage of Our Convictions*, 97 MICH. L. REV. 2381 (1999) (arguing that in order for juries to perform the important social function of taking responsibility for community decisions about individual blameworthiness and culpability, they should be aware that they are invested with this responsibility).

¹⁶⁶ Cf. Gold, *supra* note 157, at 132–33 (noting that the policies of finality and protection against juror harassment that underlie the prohibition of juror testimony to impeach a verdict must be balanced against the competing values of fairness and accuracy of verdicts).

¹⁶⁷ See Margaret A. Berger, *Expert Testimony in Criminal Proceedings: Questions Daubert Does Not Answer*, 33 SETON HALL L. REV. 1125, 1128 (2003) (“[T]he number of genetic markers now used in forensic [DNA] typing means that the likelihood that two matching profiles came from the same person or an identical twin approaches certainty, provided that the samples were properly collected and analyzed.”).

juror and witness psychology that calls into question some of the long-settled assumptions that underlie the rules of evidence.¹⁶⁸

The need to resolve disputes in a socially acceptable manner entails some method of determining the relevant historical facts to a reasonable degree of certainty. It is probably fair to say that most dispute-resolution mechanisms include a major fact-finding component, more or less accepted as reliable. In the current incarnation of the American criminal justice system, that method is the jury.¹⁶⁹ As described by Professor Nancy Marder, the conventional account of the jury is primarily grounded in this fact-finding role: “The jury’s role is narrowly envisioned by the conventional view, and to the extent juries perform more than fact finding or application of law in a narrow sense, they are seen as overstepping their bounds in the political schema and threatening the other branches’ roles.”¹⁷⁰ This conception of the jury is most apparent in judicial opinions and jury instructions,¹⁷¹ but it also underlies many academic analyses of trial proof.¹⁷² Though legal scholars most certainly tend to reject the formalistic separation of fact from law as set out in the conventional account, and tend to view the jury’s role as much more nuanced and complex than mere mechanical application of a set of clearly defined rules to the facts as found, the jury’s task of determining “what happened” in the particular case remains a crucial, if not singular, function. And rightly so: any other functions that the jury arguably carries out rest first upon some reasonably accurate account of historical fact. Fact-finding, then, is a necessary, if not sufficient, job of the adjudicatory process. And deciding whether witness testimony is credible is a significant element of adjudicative fact-finding.

¹⁶⁸ See Julie A. Seaman, *Cognitive Dissonance in the Classroom: Rationale and Rationalization in the Law of Evidence*, 50 ST. LOUIS U. L.J. 1097 (2006).

¹⁶⁹ Jury trials account for an ever smaller proportion of case dispositions in both the civil and criminal contexts. See generally George Fisher, *Plea Bargaining’s Triumph*, 109 YALE L.J. 857 (2000) (chronicling the rise of plea bargaining in criminal cases and arguing that its triumph can be explained by the interests of prosecutors and judges).

¹⁷⁰ Marder, *supra* note 26, at 905.

¹⁷¹ *Id.* at 904 (quoting representative jury instructions).

¹⁷² Professor Marder states that “few academics espouse such a view.” *Id.* She presents the conventional view of the jury in order to contrast it with a “process view,” which focuses on both “the interpretive process that goes on within the jury room as the jury deliberates” and “the process that the jury uses to communicate with other branches of government.” *Id.* at 907–08. While it is undoubtedly true that few modern academics avow the cramped and formalistic model of fact-finding seen in the conventional view as described by Professor Marder, my point here is distinct: I wish to suggest only that even the more nuanced understandings of the fact-finding and law-application process generally rest also on the jury as fact finder to some significant extent.

Yet, as Professor George Fisher has shown in his compelling historical analysis of the rise of the jury as the system's lie detector,¹⁷³ this fact-finding function has not always been viewed as belonging to the jury, even during times that the jury was vested with the task of determining guilt or punishment.¹⁷⁴ In fact, it appears that the exact opposite was the case:

[T]he lie-detecting role of juries in past generations may have been much like the lawmaking role of juries today. Today we officially declare that juries play no role in making law. Yet not only do juries manifestly make law . . . but many observers regard their power to do so as a fundamental part of our trial system. In a similar way, . . . the ideology of the jury trial system at one time required the system to *claim* that the jury did not act as lie detector, even when the jury's power to make credibility determinations was fundamental to the system's just operation.¹⁷⁵

As recounted by Professor Fisher, the modern jury's explicit role in making credibility determinations—one of its primary tasks as fact finder—slowly but steadily grew as the older means of lie detection became increasingly difficult to sustain. Just as trial by battle and by ordeal had given way to the oath as a method of determining truth,¹⁷⁶ pressures eventually built that gradually shifted ever more lie detection power from the oath to the jury.¹⁷⁷ And primary among

¹⁷³ See Fisher, *supra* note 30.

¹⁷⁴ Professor Fisher's account focuses on credibility of witnesses rather than on fact-finding in general, *see id.*, but in an age before the rise of forensic techniques, much fact-finding probably consisted of credibility determinations.

¹⁷⁵ *Id.* at 581–82.

¹⁷⁶ The medieval trials by battle and ordeal were essentially religious ceremonies that were believed to reveal truth through divine intervention. For example, in the "ordeal of the accused morsel, . . . [d]efendants were required to swallow a piece of bread after reciting a prayer that were they guilty the bread would surely choke them." VALERIE P. HANS & NEIL VIDMAR, *JUDGING THE JURY* 24 (1986) [hereinafter HANS & VIDMAR, *JUDGING THE JURY*]. "These medieval methods of proof depended on shared religious beliefs about the directing hand of God," with "the assumption . . . that God would determine which party was in the right and should prevail." VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 21–22; *see also* HANS & VIDMAR, *JUDGING THE JURY*, *supra*, at 24–25 ("The assumption underlying the ordeal, of course, was that God would intervene on the side of the innocent person."); Fisher, *supra* note 30, at 587 ("That the ordeal drew its legitimacy from the apparent intervention of God is a historical cliché, but probably true." (citations omitted)). When the Church in 1215 forbade priests from participating in trial by ordeal, "the system responded to this crisis of legitimacy by exalting a different source of divine sanction for the verdicts of its very human juries: the witness's oath, enforced (as it was thought to be) by the threat of divine vengeance." *Id.* at 589.

¹⁷⁷ Legal historian Barbara Shapiro views this shift as part and parcel of the larger intellectual movement in the seventeenth and eighteenth centuries toward empiricism and rationalism. *See* BARBARA J. SHAPIRO, *BEYOND REASONABLE DOUBT AND PROBABLE CAUSE: HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE* (1991); Barbara J. Shapiro, "To A Moral Certainty": *Theories of Knowledge and Anglo-American Juries 1600–1850*, 38 *HASTINGS L.J.* 153 (1986) [hereinafter Shapiro, *Moral Certainty*]; Barbara J.

these pressures was a growing sense, which the system strove to submerge with various “contorted maneuverings,” that the system’s reliance on the oath as a guarantor of truth was questionable.¹⁷⁸

The various rules that worked to protect the legitimacy of the oath as a means of ensuring truthful testimony¹⁷⁹ reveal some interesting similarities to current evidentiary rules surrounding the jury’s fact-finding role. The first line of defense employed to protect the power of the oath was to prevent conflicting oaths from being heard.¹⁸⁰ Thus, criminal defendants early on were prevented from calling witnesses at all.¹⁸¹ Similarly, as noted above, the first response of courts to evidence that might call into question the jury’s ability to determine credibility has generally been to bar such evidence altogether, citing its potential to “usurp the role of the jury.”¹⁸² Modern courts have generally been as solicitous of the jury’s fact-finding role as they have been suspicious of its law-finding role. Though in some areas this protection of the jury’s traditional fact-finding function has sounded more in rhetoric than in practical operation,¹⁸³ when it comes to evidence as to the veracity of witnesses—whether via expert or machine—courts have tended to place an extremely high bar on admissibility.¹⁸⁴

Shapiro, *Law and Science in Seventeenth-Century England*, 21 STAN. L. REV. 727 (1969) [hereinafter Shapiro, *Law and Science*].

¹⁷⁸ Fisher, *supra* note 30, at 704. Fisher writes:

As we now know, these contorted maneuverings to protect the system’s legitimacy actually eroded that legitimacy. Bentham aimed his most withering ridicule at the old system of competency rules, which he said purported to guard the truth but in fact merely hid from the jury many potential sources of truth. But long before Bentham and long before the Enlightenment or the Scientific Revolution, ordinary people must have shaken their heads in wonderment at a system designed to find truth that permitted only one party to the cause to present any sworn evidence at all.

Id. at 704–05.

¹⁷⁹ “With an elaborate array of evidentiary devices, the system managed to preserve for a very long time the principle that guilty verdicts in capital cases (and most verdicts in most other cases) should never rest on a merely human choice between two competing versions of a divinely revealed truth.” *Id.* at 602.

¹⁸⁰ *Id.* at 580.

¹⁸¹ *Id.* at 603.

¹⁸² *Id.* at 601.

¹⁸³ In the civil context especially, the rising use of procedures such as summary judgment, special verdicts, and remittitur has weakened the power of the jury relative to the judge. See generally Thomas, *The Seventh Amendment*, *supra* note 128; Thomas, *Judicial Modesty*, *supra* note 128.

¹⁸⁴ The admissibility of scientific evidence in the federal courts is governed by Federal Rule of Evidence 702, which provides as follows:

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert . . . may testify

Over time, as the rules that simply prevented defendants from calling witnesses at all yielded to various social, political, and intellectual developments,¹⁸⁵ there evolved other rules that continued to reinforce the illusion that the oath served a significant lie detection function. These included “a broad series of witness competency rules that barred whole categories of witnesses—those thought most likely to lie—from testifying,” and an instruction that “advised juries to reconcile conflicting testimony in a way that would avoid branding any witness a perjurer.”¹⁸⁶ These rules, in turn, gave way to others that required corroboration for certain categories of testimonial proof, and to formalistic instructions regarding how the jury was to resolve conflicting testimony. One can view this trajectory as the system’s attempt to manage the admissibility of evidence that appeared increasingly necessary to an accurate assessment of the adjudicative facts, yet which challenged the reigning paradigm of fact-finding.¹⁸⁷

In the modern context, the main challenge thus far to the jury’s primacy as fact finder—and especially as lie detector—has come in the guise of scientific and expert evidence that bears on the credibility of witnesses. As was the case

thereto in the form of an opinion or otherwise if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

FED. R. EVID. 702.

In the trilogy of cases interpreting and applying Rule 702, beginning with *Daubert* in 1993, the Supreme Court has held that the rule imposes a heightened reliability requirement upon the admissibility of scientific (and other expert) evidence, and that trial courts are to act as “gatekeepers” to ensure that only reliable expert evidence reaches the jury. See *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 148 (1999); *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 142 (1997); *Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579, 597 (1993). Even given the heightened scrutiny used to screen expertise, courts are exceptionally wary of evidence that tends to bear on the credibility of other witnesses in the case. See Pager, *supra* note 131, at 375. This extra wariness is not merely a post-*Daubert* development. See Adria Goodkin Kaplan, *The Lie Detector: An Analysis of Its Place in the Law of Evidence*, 10 WAYNE L. REV. 381, 381 (1964) (comparing judicial treatment of polygraph evidence to other scientific evidence, decades before *Daubert*, and concluding that “courts, in considering the admissibility of lie detector evidence, have not merely excluded the evidence but have judged it by a different standard than the standard which is established for determining admissibility and applied to other scientific evidence”).

¹⁸⁵ See Fisher, *supra* note 30; Shapiro, *Moral Certainty*, *supra* note 177, at 153–55 (arguing that the development of the concept of “moral certainty” in jury fact-finding was linked to the broader intellectual currents of philosophy and science); Shapiro, *Law and Science*, *supra* note 177, at 727–30 (arguing that the scientific revolution in seventeenth-century England was inextricably entwined with the development of legal epistemological thought and its expression in beliefs about jury determination of truth).

¹⁸⁶ Fisher, *supra* note 30, at 584.

¹⁸⁷ See SHAPIRO, *supra* note 177, at 244 (“[T]he insistence on the old ways was not mere anachronism. It provided a legitimating screen that protected the jury’s old claim to truth-telling [based on their own personal knowledge] while a new one was sought.”).

with challenges to the primacy of the oath, the initial inclination to close the courtroom doors to such evidence has gradually given way to more complicated rules that allow some such evidence, while containing its impact upon the legitimacy of the jury as lie detector. For instance, expert testimony regarding the credibility of child witnesses is restricted such that an expert is not permitted to opine on the credibility of a specific witness, but “this has not resulted in the wholesale exclusion of experts on credibility[:]; it has only affected the scope of what they might comment upon.”¹⁸⁸ Scientific evidence that directly addresses witness credibility has generally engendered very strong resistance by courts, but as its relevance and reliability have seemed increasingly arguable,¹⁸⁹ the system has begun to admit it, subject to various constraints.¹⁹⁰ With respect to polygraph evidence, one common compromise is to admit the evidence only upon stipulation;¹⁹¹ likewise, the earliest jury trials were held only upon consent of the accused.¹⁹²

¹⁸⁸ See DAVID L. FAIGMAN ET AL., SCIENCE IN THE LAW: SOCIAL AND BEHAVIORAL SCIENCE ISSUES 323 (2002) (describing a number of ways in which such credibility evidence is admitted despite the general ban).

¹⁸⁹ While an evaluation of the reliability of polygraphy is certainly beyond the scope of this Article, some courts have recently deemed it sufficiently reliable to be admissible under limited circumstances, although there is much debate among experts as to its reliability. One thing that can be said is that it is at least as reliable as several categories of forensic evidence that have routinely been admitted against criminal defendants. See generally DAVID L. FAIGMAN ET AL., SCIENCE IN THE LAW: FORENSIC SCIENCE ISSUES (2002) (detailing reliability concerns raised by a number of forensic techniques). Other types of expert evidence relating to reliability of eyewitness identification are well-supported in the relevant social science literature. FAIGMAN ET AL., *supra* note 188, at 419 (surveying the research and stating that “issues of disagreement among eyewitness experts pertain less to disagreement over the scientific findings in the research and more to the question of how these findings can be generalized to actual cases facing the courts and the question of whether expert opinion testimony is beneficial to the trier of fact”).

¹⁹⁰ See, e.g., *United States v. Belyea*, 159 Fed. App’x 525, 529 (4th Cir. 2005) (permitting testimony of expert on false confession, because “[t]he phenomenon of false confessions is counter-intuitive and is not necessarily explained by the general proposition that ‘jurors know people lie’”); *United States v. Gonzalez-Maldonado*, 115 F.3d 9, 16–17 (1st Cir. 1997) (concluding that trial court erred by refusing to allow psychiatrist to testify that defendant, as a result of mental illness, symptomatically exaggerated as it could be relevant to the “credibility of [taped] statements”); *State v. St. Germain*, 153 P.3d 591, 597 (Mont. 2007) (“Expert testimony is admissible . . . for the purpose of helping the jury to assess the credibility of a child sexual assault complainant under certain circumstances.”); *Darling v. State*, 262 S.W.3d 920, 924 (Tex. Ct. App. 2008) (“A trial court does not err when it admits an expert witness’ opinion testimony that a child did not exhibit indications of having been coached to make a false accusation of sexual abuse.”).

¹⁹¹ See, e.g., *United States v. Piccinonna*, 885 F.2d 1529, 1535 (11th Cir. 1989) (holding that polygraph evidence is not per se inadmissible, as parties can stipulate to admit polygraph evidence); *United States v. Gordon*, 688 F.2d 42, 44 (8th Cir. 1982) (holding that polygraph evidence is admissible if parties stipulate in advance of test). But see *United States v. A&S Council Co.*, 947 F.2d 1128, 1134 (4th Cir. 1991) (declining to depart from a per se rule banning polygraph evidence).

¹⁹² Of course, the accused was often persuaded to “consent” by the alternative, known as the *peine forte et dure* (“strong and continuing pain”), whereby he was made to lie under heavy stones until he either consented to be tried by a jury or was crushed to death. See VIDMAR & HANS, AMERICAN JURIES, *supra* note 20, at 24–35.

The antecedents to the jury as fact finder and lie detector ultimately gave way to pressures that undermined their legitimacy by undermining their reliability as methods for determining historical truth.¹⁹³ The jury has thus far proved to be “a remarkably reliable source of systemic legitimacy,”¹⁹⁴ but its reliability as a lie detector and finder of adjudicative fact is currently under increasing pressure from several sources. Most glaring among these is the steady flow of media stories about post-conviction exoneration of innocents,¹⁹⁵ including several who had been sentenced to death.¹⁹⁶ Particularly in those cases in which the exonerations were accomplished through DNA testing of retained evidence, it is almost certain that juries convicted defendants who

¹⁹³ See SHAPIRO, *supra* note 187, at 3 (“A much deeper [intellectual] crisis occurred during the twelfth century, when ‘irrational proofs,’ such as trial by ordeal, could no longer be seen as consistent with justice or with how truth determinations ought to be made.”). Though truth and legitimacy are certainly distinct functions of the jury trial, legitimacy is closely tied to the system’s ability (real or apparent) to discover the truth. Whether protecting legitimacy at the expense of truth, for example, by hiding flaws in the system, is either possible or desirable is a separate question.

¹⁹⁴ Fisher, *supra* note 30, at 705.

¹⁹⁵ There have been hundreds of such stories the last ten years. The *New York Times* recently ran a front-page story marking the two hundredth DNA exoneration, complete with profiles and photographs of several of the exonerees and an interactive web-based feature profiling nearly all of the others. See Chris Conway, *The DNA 200*, N.Y. TIMES, May 20, 2007, § 4, at 14 (chronicling the stories of 200 inmates formally cleared on the strength of DNA evidence); see also Kevin Johnson, *DNA Tests Fuel Urgency to Free the Innocent*, USA TODAY, Feb. 19, 2008, at 1A (“After spending nearly 27 years buried in the vast Texas prison system for a crime he did not commit, Charles Chatman’s first weeks of freedom have been overwhelming.”). Other popular media have run similar stories. See, e.g., *60 Minutes: The D.A. and the Death Penalty* (CBS television broadcast May 4, 2008) (documenting the exoneration of James Woodward by DNA evidence after serving twenty-seven years in prison for the murder of his girlfriend); *NBC Nightly News: Who We Are: The Truth About DNA* (NBC television broadcast Feb. 12, 2008) (highlighting the wrongful conviction and subsequent exoneration by DNA evidence of Kennedy Brewer); *Primetime: Man Spends 17 Years in Jail Based on Apparent False Confession* (ABC television broadcast Mar. 29, 2006) (showcasing how DNA evidence cleared three men after they each served seventeen years in prison for the rape and murder of a young girl). In addition, a recent best-selling nonfiction book by author John Grisham tells the story of the exoneration of a man who had been convicted of murder and sentenced to death. JOHN GRISHAM, *THE INNOCENT MAN: MURDER AND INJUSTICE IN A SMALL TOWN* 308–12 (2006). Several popular movies and documentaries have also depicted the theme of wrongful conviction. See *THE THIN BLUE LINE* (Miramax Films 1988) (depicting the wrongful conviction of a man accused of the murder of a police officer); *THE HURRICANE* (Universal Pictures 1999) (chronicling the story of a man wrongfully convicted of murder and his subsequent exoneration after nineteen years in jail); *AFTER INNOCENCE* (American Film Foundation 2005) (examining the cases of seven men wrongly convicted of murder and rape and exonerated years later by DNA evidence).

¹⁹⁶ See Lawrence C. Marshall, *Litigating in the Shadow of Innocence*, 68 U. PITT. L. REV. 191, 195 (2006) (book review) (“[Although] there are those who challenge whether *some* of those often identified as wrongly convicted actually committed the crime for which they had once been convicted . . . [this] does not detract from the fact that even the most severe critics of the lists *agree that dozens of actually-innocent defendants have been sentenced to death in the past decades.*”).

were factually innocent.¹⁹⁷ Though the scholarly literature is somewhat mixed regarding the systemic implications of these exonerations,¹⁹⁸ the public impression left in the wake of such media accounts must certainly be one of skepticism about the system's infallibility.¹⁹⁹ Indeed, regardless of the actual incidence of false conviction, the public perception of false conviction affects the system's legitimacy.²⁰⁰ After some initial intransigence,²⁰¹ several legislatures, and even some prosecutors and other executive officials, have responded to the exoneration accounts with proposals for conviction reviews as well as reforms aimed at preventing those factual errors that appear most likely to contribute to wrongful convictions.²⁰² These early efforts may prove successful—in increasing the accuracy of the system as well as bolstering the public perception of its accuracy—but in the meantime, the steady diet of

¹⁹⁷ As of November 16, 2008, there had been 223 DNA exonerations. See The Innocence Project, <http://www.innocenceproject.org> (last visited Nov. 16, 2008).

¹⁹⁸ See sources cited *supra* note 31.

¹⁹⁹ See VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 16 (“Over the last decade, cases in which DNA evidence has exonerated defendants who had been convicted by juries—including some prisoners on death row—have raised questions about how well juries evaluate trial evidence and reach their verdicts.”).

²⁰⁰ According to one recent poll, “94% of Americans believe innocent people are sometimes convicted of murder; those polled estimated that this happens 13% of the time.” Pager, *supra* note 131, at 411 (citing Humphrey Taylor, *The Harris Poll*, #41 (Aug. 2, 2000), http://www.harrisinteractive.com/harris_poll/index.asp?PID=101).

²⁰¹ See Susan Bandes, *Loyalty to One's Convictions: The Prosecutor and Tunnel Vision*, 49 *How. L.J.* 475, 475 (2006) (recounting and offering psychological explanation for the persistent phenomenon of “the refusal of prosecutors to concede that the wrong person was convicted, even after a defendant's exoneration”); Daniel S. Medwed, *The Zeal Deal: Prosecutorial Resistance to Post-Conviction Claims of Innocence*, 84 *B.U. L. REV.* 125, 125–28 (2004); see also *FLA. R. CRIM. P.* 3.853(d) (repealed 2006) (imposing a two-year statute of limitations on applications for post-conviction DNA testing). The judiciary, too, has tended to be markedly unsympathetic to post-conviction claims based on factual innocence. Jerome M. Maiatico, Note, *All Eyes on Us: A Comparative Critique of the North Carolina Innocence Inquiry Commission*, 56 *DUKE L.J.* 1345, 1352–53 (2007) (“Federal law makes it difficult for claims of factual innocence to be heard by federal courts.”). One certainly could, if so inclined, construct a rather convincing story of a system intent on burying evidence of wrongful conviction wherever possible. Cf. Fisher, *supra* note 30, at 579–81 (discussing the historical account of the use of the oath as the system's lie detector, and suggesting that the primary goal of the system was to protect its own legitimacy by preventing instances in which the reliability of the oath as a guarantor of truthfulness would be called into question).

²⁰² See H.R. 1071, 149th Gen. Assem., Reg. Sess. (Ga. 2008) (commonly referred to as the Witness Identification Accuracy Enhancement Act; calling for the Georgia Peace Officer and Training Council to establish guidelines and procedures for conducting show-ups and line-ups); H.B. 352, 2007 Gen. Assem., Reg. Sess. (N.C. 2007) (mandating blind administration, specific instructions to witnesses, training, and remedies for non-compliance for consideration by the court); Ralph Blumenthal, *For Dallas, New Prosecutor Means an End to the Old Ways*, *N.Y. TIMES*, June 3, 2007, at N28 (discussing Dallas District Attorney Craig Watkins's order to review post-conviction DNA evidence on more than 400 cases); *CNN: Illinois Suspends Death Penalty* (CNN television broadcast Jan. 31, 2000) (detailing former Illinois governor George Ryan's moratorium on the state's death penalty and investigation into finding out why more executions have been overturned than carried out since 1977 in Illinois).

stories about wrongful convictions creates a stress upon the legitimacy of the jury as a fact finder and a lie detector.

At the same time, the ability of juries to determine certain kinds of adjudicative facts has been increasingly called into question by other developments. For example, there is widespread skepticism about jurors' ability to understand and evaluate the increasingly prevalent and often complex scientific evidence presented in many trials.²⁰³ With the advent and rapid advancement of DNA technology, there is a danger that juries will be viewed as virtually dispensable in cases in which there is overwhelming scientific evidence of guilt. In addition, the newfound ability of the public essentially to sit as the thirteenth juror, through such media channels as truTV (formerly known as CourtTV), now allows viewers to evaluate the evidence along with the jury. As viewers inevitably disagree with jury verdicts in such cases,²⁰⁴ they may come to doubt juries' ability to determine facts accurately.²⁰⁵

Each of these developments exerts some slight pressure on the legitimacy of the jury system as a reliable fact finder,²⁰⁶ and each is likely to lead to some efforts at reform, around the margins, aimed at increasing the reliability of jury fact-finding.²⁰⁷ Suppose, however, that it were to become apparent that a

²⁰³ See VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 169–89 (summarizing this critique of jurors' abilities and evaluating its validity in light of the available empirical evidence); Joseph Sanders, *The Merits of the Paternalistic Justification for Restrictions on the Admissibility of Expert Evidence*, 33 SETON HALL L. REV. 881, 891–99 (2003) (assessing the empirical research on the ability of jurors to assess scientific evidence); Seaman, *supra* note 44, at 869–76 (discussing distrust of juries' ability to evaluate expert opinion testimony).

²⁰⁴ See Ruprecht, *supra* note 164, at 245 (“[T]here have been recent cases where a large segment of the national population has observed the trial and evidence almost as closely as the jury, and there has been widespread disagreement with the verdicts rendered.”).

²⁰⁵ A written account or artist's depiction cannot convey the full measure of testimony or other evidence, so there is always the chance that the outside viewer disagrees because she has missed some nuance that the jury has observed. This possibility is reinforced by the notion—almost certainly false—that demeanor evidence is a reliable indicator of truthfulness. See, e.g., Wellborn, *supra* note 36, at 1075–76 (describing studies demonstrating that demeanor is a poor indication of truthfulness).

²⁰⁶ Of course, the competency of the jury in this regard must be measured against the relative competency of other methods of determining adjudicative fact. Most research demonstrates that juries compare favorably in this regard with judges. See VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 148–51 (describing Kalven and Zeisel's classic study of juries that found substantial rates of judicial agreement with jury verdicts, as based on juries' competency and fact-finding abilities).

²⁰⁷ See, e.g., Giannelli, *supra* note 31. Note, however, that there have been reform proposals aimed at improving the accuracy of jury fact-finding and law application grounded upon strong empirical research for years, but the judiciary seems strangely resistant to adopting them. See Michael J. Saks, *Judicial Nullification*, 68 IND. L.J. 1281, 1289 (1993) (suggesting that by offering juries instructions that cannot be understood, judges regularly preserve an element of jury nullification). Legislatures have been somewhat more receptive to such reforms.

technology existed that could truly resolve factual disputes; when combined with the specter of actual innocence, could—and should—the system exclude such evidence in the service of preserving the jury’s traditional role as sole judge of the facts? Should the system exclude such evidence in order to keep the black box tightly shut?

C. *The Box Splits Open: What Remains of the Jury’s Role?*

Some who have addressed these questions have argued that, even in the face of an extremely accurate neuroscience-based lie detector, a jury role would still remain. First and most obviously, with respect to lay witness testimony, not all subjective truths are actually accurate statements about the world. A witness might believe that she is testifying truthfully but be mistaken or perhaps delusional;²⁰⁸ it is sometimes assumed that even a foolproof lie detector would not detect such untruths.²⁰⁹ Second, with regard to expert witness testimony—including expert testimony on the subject of lie detection and whether another witness’s testimony evidenced neural correlates of deception—there would remain the usual jury issues as to the credibility of the expert: bias, defects in test methodology, the reliability of the particular machinery used, or even outright corruption and deceit.²¹⁰ Third, even assuming an accurate recounting of events—in the sense of who did what to whom—there would still in many cases remain the “why question” of the intent of the relevant actor.²¹¹ Finally, constitutional protections or ethical

²⁰⁸ Spence et al., *Munchausen’s Syndrome*, *supra* note 17, at 311–13.

²⁰⁹ Cf. Jerome Frank, *A Plea for Lawyer-Schools*, 56 *YALE L.J.* 1303, 1307 (1947) (“We have not yet perfected a foolproof lie detector; we certainly now have no detector of the unconscious distortions of prejudiced but honest witnesses; and almost surely we will never have a contrivance for correcting a witness’ original mistaken observation of the facts.”); Helen Pearson, *Lure of Lie Detectors Spooks Ethicists*, 441 *NATURE* 918, 919 (2006) (“[D]ata collected from healthy subjects reveal little about the mindset of someone who genuinely believes they are telling the truth or someone who is confused, delusional or a pathological liar.”).

²¹⁰ See Pardo, *supra* note 27, at 318 (noting that, as with other types of scientific expert testimony, the jury would still need to “consider whether other evidence regarding credibility should override the test results, rendering the test conclusion unlikely . . . [due to] the possibility of errors in conducting or analyzing the test, . . . [or] the possibility of perjury by the technician”).

²¹¹ See Spence & Kaylor-Hughes, *Looking for Truth and Finding Lies*, *supra* note 33, at 68 (“[W]hile a body might undoubtedly leave evidence that ‘it’ was present at the scene of a crime, and while film might recapitulate the movements which that body ‘made’, we are not simply logging the motions of machines here. . . . [W]e are crucially concerned with intentions, states of mind, with what the subject was thinking.”). Of course, witnesses other than the accused could be asked about their relevant states of mind, and the accused could be examined should he elect to testify. However, even in such cases the issue of intent is often much less a question of historical fact than one of normative judgment. See Hoffman, *supra* note 31, at 663 (“The relatively few felony cases that actually go to trial in America are typically about moral guilt, not factual

considerations might prevent the government or private parties from looking inside a criminal defendant's mind or from coercing other individuals to undergo neuropsychological lie detection procedures.²¹²

Each of these various observations about what would remain even in the face of a very reliable lie detection technology tends to highlight a distinct aspect of the jury's function. The first point goes to the jury's conventional fact-finding role: The argument that neuroimaging lie detection technology would not detect falsehoods that the individual subjectively believes to be true, thus leaving the jury's role of determining witness credibility intact, is an empirical prediction about the limits of the technology.²¹³ Certainly many—perhaps most—objectively untrue statements uttered by trial witnesses are mistakes, exaggerations, or distortions rather than bald-faced intentional lies.²¹⁴ Yet it does not necessarily follow that these will forever be immune to detection through fMRI or other as-yet-unknown technologies. Even now, scientists have had some success in distinguishing true from false memories based upon differences in brain activation.²¹⁵ It appears that individuals may implicitly recognize things they have observed, even when they have no conscious awareness of having observed them.²¹⁶ At the molecular level, there

guilt.”); see also H. Richard Uviller, *Acquitting the Guilty: Two Case Studies on Jury Misgivings and the Misunderstood Standard of Proof*, 2 CRIM. L.F. 1, 19–20 (1991) (discussing example of an “inherently nonsensical” jury instruction on the distinction between criminal mental states for murder).

²¹² See Pardo, *supra* note 27, at 321–37 (analyzing the constitutionality of neuroscience-based lie detection under the Fourth and Fifth Amendments and the Due Process Clause). I informally poll my evidence students on this question every semester and have done so at least five times; results consistently show a significant minority of students who would be uncomfortable with the prospect of an infallible courtroom lie detector based on notions of “neural privacy.”

²¹³ Discussing the conceptual and empirical difficulties posed by delusional patients, Professor Langleben states that “[t]heoretically, delusions and other impairments in reality testing that are indistinguishable from the truth on the cognitive level could be inaccessible to either the polygraph or brain-imaging based lie-detection. Experimental data from well-controlled experiments on psychophysiological lie-detection in delusional subjects are required to confirm or refute this assumption.” Langleben et al., *True Lies*, *supra* note 72, at 363.

²¹⁴ See, e.g., Elizabeth F. Loftus, *Make-Believe Memories*, 58 AM. PSYCHOLOGIST 864, 867–69 (2003) (canvassing two decades of research on the dangers of eyewitness inaccuracy); Gary L. Wells & Elizabeth A. Olson, *Eyewitness Identification*, 54 ANN. REV. PSYCHOL. 277, 277–78 (2003).

²¹⁵ See Nobuhito Abe et al., *Neural Correlates of True Memory, False Memory, and Deception, Cerebral Cortex*, 18 CEREBRAL CORTEX 2811 (2008) (finding differences in brain activation between deception and false memory); Hongkeun Kim & Robert Cabeza, *Trusting Our Memories: Dissociating the Neural Correlates of Confidence in Veridical Versus Illusory Memories*, 27 J. NEUROSCI. 12,190 (2007) (fMRI study finding that distinct brain regions associated with high-confidence recollection of true memories and high-confidence illusory (false) memories).

²¹⁶ See, e.g., Kristina Schutz et al., *Nice Wor_ if You Can Get the Wor_ : Subliminal Semantic and Form Priming in Fragment Completion*, 16 CONSCIOUSNESS & COGNITION 520 (2007).

have been significant advances in scientific understanding of how individual neurons encode knowledge.²¹⁷ In sum, insofar as it is conceivable that neuroscience will one day enable accurate lie detection, it seems equally conceivable that it could allow detection of other kinds of untruths that are not, strictly speaking, deceptive in the sense of intentionally misleading.²¹⁸

The remaining points, in contrast, implicate jury functions that are distinct from that of determining historical, adjudicatory fact. Questions about the integrity of the lie detection expert and method, the intent of the relevant actors, and potential constitutional or ethical constraints on the use of the technologies call upon the jury to perform its more hidden roles. Rather than touching on the jury's responsibility to determine the historical facts, these latter points highlight the jury's less remarked role, broadly conceived, of interpreting and applying the law.²¹⁹

²¹⁷ See Underwood, *supra* note 140, at 623–33.

²¹⁸ Furthermore, even a lie-detector that could only distinguish subjective deception would offer valuable evidence to the jury, since individuals commonly err in lie detection by classifying truth-tellers as deceptive. See VRII, *supra* note 33, at 344–57 (citing studies on lay errors). *But see* Pardo, *supra* note 27, at 315 (arguing that the brain state is not identical with the “lie” or “knowledge”).

²¹⁹ In a series of classic articles, Professor Kenneth Culp Davis suggested the terms “adjudicative fact” and “legislative fact” to denote a distinction between two types of material relied upon by courts in judicial decision making. See Davis, *Judicial Notice I*, *supra* note 38, at 952–59; Davis, *Judicial Notice II*, *supra* note 38, at 524–28; Davis, *An Approach to Problems*, *supra* note 38, at 404–07. Though Professor Davis was concerned with rules governing judicial notice and with facts as used by courts and administrative agencies, his description of the distinction is also apt in the context of jury function:

When a court or an agency finds facts concerning the immediate parties—who did what, where, when, how, and with what motive or intent—the court or agency is performing an adjudicative function, and the facts so determined are conveniently called adjudicative facts. When a court or an agency develops law or policy, it is acting legislatively: the courts have created the common law through judicial legislation, and the facts which inform the tribunal's legislative judgment are called legislative facts.

Davis, *Judicial Notice I*, *supra* note 38, at 952.

Similarly, when the jury performs its classic fact-finding function of determining what happened in the case being tried, it is performing the adjudicative function described by Professor Davis. In contrast, when it performs other functions such as determining moral blameworthiness, checking the conduct of the prosecutorial authority, or evaluating the justice of the application of a law to the adjudicative facts, it is performing a legislative function and the facts it consults are “extra-record facts,” more aptly termed legislative facts. As some scholars have pointed out, nullification and law-finding have often been conflated though they are in fact distinct, *see* Akhil Reed Amar, *The Bill of Rights as a Constitution*, 100 YALE L.J. 1131, 1191 n.264 (1991); David A. Pepper, *Nullifying History: Modern-Day Misuse of the Right to Decide the Law*, 50 CASE W. RES. L. REV. 599, 601 (2000); however, for purposes of this discussion I refer to “law-finding” as everything the jury does that is not finding historical fact.

Facts surrounding the expert presentation of the (assumed reliable) lie detection evidence primarily implicate the jury's institutional character as a check upon the power of the state vis-à-vis the individual.

Questions of intent require the jury to apply its sense of moral blameworthiness. As Professor Michael Risinger has noted:

States of mind are not factual in the same way that determining it was cloudy at noon in Newark last Wednesday is factual, despite the rather flippant but oft-cited observation of Lord Justice Bowen that "the state of a man's mind is as much a fact as the state of his digestion." With all due respect to Lord Justice Bowen, it just ain't so.²²⁰

Finally, extrinsic constitutional or ethical considerations most obviously draw the jury beyond its conventional fact-finding role to act as a bulwark against governmental intrusion or overreaching.²²¹ Thus, the jury functions that arguably would remain even in the face of an extremely accurate lie detector would be precisely those functions that are currently least obvious and most contested.²²² That is, to the extent that the historical facts are given, the task of the jury necessarily shifts toward deciding whether and how to apply the law (as received through the court's instructions) to those facts. And if the fact-finding task were effectively removed from the mix of jury responsibilities because it could be accomplished by a reliable machine, the exercise of this quasi-legislative power would become virtually unmistakable.

III. LEGISLATING INSIDE THE BLACK BOX: JURY NULLIFICATION

Distrust of the jury is not confined to doubts about its ability accurately to determine historical facts. Rather, the twin strands of distrust of the jury correspond roughly to its two general functions of fact-finding and

²²⁰ Risinger, *supra* note 38, at 1293.

²²¹ To be sure, constitutional criminal procedural constraints are applied by the courts, which may therefore preempt or override their being determined by the jury. However, the jury continues to provide a check in the sense that it could decide to acquit a defendant where it determines that the evidence was wrongfully procured even though a court has deemed it constitutionally permissible. *See, e.g.*, Darryl K. Brown, *Jury Nullification Within the Rule of Law*, 81 MINN. L. REV. 1149, 1172-73, 1172 n.100 (1997) (recounting actual and hypothetical cases in which a court admits unlawfully seized evidence by accepting the patently perjured account of a police officer, but the jury is suspicious of the account, suspects government wrongdoing, and therefore acquits the defendant). In a civil case, it might similarly discount or disregard evidence that, in its judgment, is ethically tainted.

²²² These remaining functions loosely map onto the types of jury nullification described by a leading nullification scholar. *Id.* at 1171-96.

lawmaking;²²³ to adapt Professor Kenneth Culp Davis's terminology, they correspond to the jury's adjudicative and legislative functions.²²⁴ As discussed in Part II, doubts about the jury's adjudicative ability have been bolstered in recent years due to highly publicized DNA exonerations that have shone a spotlight on failures in individual cases, and also to a growing body of research that calls into question the accuracy of eyewitness identification and the human ability to detect lies.²²⁵ In addition, the increasing importance of complex scientific evidence to a wide range of cases has further undermined public confidence in the jury's ability to make accurate fact determinations. These developments should increase pressure on the system to adopt reliable methods, if they exist, that can potentially aid the jury in its fact-finding function. At the same time, such technologies will tend to decrease the work that, according to conventional wisdom, is the jury's singular function.

With respect to the jury's legislative function, however, distrust is if anything even more pronounced than distrust of its fact-finding capabilities. That distrust reaches its clearest expression in the debate over the jury's power to "nullify" the law, because nullification is the purest example of a jury's exercise of a legislative function.

While the jury as an institution has always engaged in some form of lawmaking, that role has sharply diminished over time.²²⁶ It remains most salient today in the paradigmatic example of jury nullification—a criminal case in which jurors decide to "acquit the defendant even though they understand that the defendant is guilty under the law and the facts."²²⁷ Indeed, many

²²³ As Vidmar and Hans write, "over the past several decades American criminal and civil juries have been criticized for incompetence and irresponsibility." VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 15. "Incompetence" goes to distrust of the jury's adjudicative ability, while "irresponsibility" suggests distrust of its legislative role.

²²⁴ See Davis, *An Approach to Problems*, *supra* note 38, at 404–07.

²²⁵ *But see* Ruprecht, *supra* note 164, at 245 n.119 (quoting noted jury expert Harry Kalven's 1955 memo to Congress, in which Kalven wrote, "Nor, contrary to my first impression, does there seem to be any particular period in which the debate [over the jury system] grows hotter or colder. It has always been a hot debate.>").

²²⁶ See generally CONRAD, *supra* note 40; VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 54–64. Even among critics of nullification, the account of the historical power of English, colonial, and early American juries to decide the law as well as the facts is accepted. See Pepper, *supra* note 219, at 608–09 ("Undoubtedly, the early American experience included the jury's right to decide the law. As nullification scholars point out, evidence abounds of that right: statements by the Founders, cases and judicial statements at all levels, tracts, state constitutions, statutes, and other sources." (citations omitted)).

²²⁷ VIDMAR & HANS, *AMERICAN JURIES*, *supra* note 20, at 227. Though the term is often used imprecisely, most scholars define nullification to require an intent on the part of the jury to reject the court's legal instruction (as opposed, for example, to confusion or misunderstanding of the court's instruction). *E.g.*,

commentators appear to view jury nullification as the last remnant of the jury's lawmaking function,²²⁸ akin to vestigial human traits like the appendix or tailbone. And just as there is debate over the continued usefulness of such physical structures, there is much debate about the reason for, and value of, jury nullification: Is it a necessary evil that must occasionally be tolerated, but that should be constrained insofar as possible, or a desirable—indeed crucial—aspect of the jury as an institution? Is it a power to be hidden from the jury, or a right about which the jury should be instructed in every case? Is it primarily an opportunity for the jury to exercise its illegitimate social biases, or to protect individual liberty as against excesses of state power?

The controversy over jury nullification is a crystallization of the more general ambivalence with which the American jury is regarded.²²⁹ It is typically said that when a jury nullifies, it implements its own idiosyncratic view of justice or morality²³⁰—either in the individual case or more

Conrad, *supra* note 39, at 11; Hannaford-Agor & Hans, *supra* note 45, at 1254; Marder, *supra* note 26, at 881. In addition, most scholars confine their definition and analyses to criminal cases because of the much more robust power of the court in civil cases to overturn jury verdicts that are against the weight of the evidence and to control jury action through the use of special verdicts. *E.g.*, CONRAD, *supra* note 40, at 7. *But see* Lars Noah, *Civil Jury Nullification*, 86 IOWA L. REV. 1601, 1603 (2001) (arguing that a definition of jury nullification that excludes civil cases is “unduly narrow” because “trial and appellate judges often hesitate to intrude on the jury’s fact-finding domain,” and thus civil “verdicts based on a jury’s intentional deviation from the law as instructed by the court may escape post-trial correction”); Kaimipono David Wenger & David A. Hoffman, *Nullificatory Juries*, 2003 WIS. L. REV. 1115 (likening jury role in punitive damages context to criminal jury nullification). Finally, most focus on acquittals against the evidence, because nullificatory convictions are subject to correction by the court. *E.g.*, Hannaford-Agor & Hans, *supra* note 45, at 1255.

²²⁸ To be sure, some critics of jury nullification view such acts as “lawless”—outside the law—and therefore it might be misleading to characterize their view of what a nullifying jury is doing as “lawmaking.” Most, however, take primary issue with jury nullification based upon a critique of the process rather than from a categorical argument that such a decision does not amount to some form of lawmaking. In contrast, David Pepper has argued that there is a fundamental distinction between interpreting law on the one hand (which juries historically were empowered to do), and rejecting law-as-interpreted on the other. Thus, according to this argument, the logical leap of the pro-nullification scholars from the historical power of juries to the legitimacy of nullification is flawed. *See* Pepper, *supra* note 219, at 608–09; *see also* Amar, *supra* note 219, at 1191 (distinguishing between question of jury power to review laws for constitutionality and power to disregard laws regardless of constitutionality).

²²⁹ *See* CONRAD, *supra* note 40, at 4 (noting that although “Americans maintain a practically religious devotion to the institution of trial by jury, we remain ambivalent about what juries in criminal cases are supposed to do”); Albert W. Alschuler, *The Supreme Court and the Jury: Voir Dire, Peremptory Challenges, and the Review of Jury Verdicts*, 56 U. CHI. L. REV. 153, 153–54 (1989) (describing ambivalent views of the criminal jury held by defense attorneys, prosecutors, and judges); Mark Cammack, *The Jurisprudence of Jury Trials: The No Impeachment Rule and the Conditions for Legitimate Legal Decisionmaking*, 64 U. COLO. L. REV. 57, 57 (1993) (“There is probably no other feature of the law that is at once the object of such extravagant praise and utter contempt as the lay jury.”); Seaman, *supra* note 44, at 869–76.

²³⁰ *E.g.*, Crispo et al., *supra* note 39, at 3 (“[T]he nullifying jury serves as a mini-legislature, repealing laws it deems unjust or preventing what it sees as the harsh, inequitable application of law in certain cases.”).

generally.²³¹ In the individual case, nullification permits the jury to tailor the law to the particular situation, so that a just law is not applied in an unjust manner.²³² Where nullification is more general, it demonstrates a mismatch between the law and community sentiment.²³³ In either case, the jury is substituting its own idea of what is moral or just for that arrived at through the usual democratic processes.²³⁴ For this reason, nullification is often portrayed by its critics as an example of lawlessness, usurpation, and anarchy.²³⁵

At least since the Supreme Court's landmark 1895 decision in *Sparf & Hansen v. United States*,²³⁶ the federal courts have almost uniformly treated jury nullification as illegitimate.²³⁷ *Sparf*, like most of the nullification cases

²³¹ Professor Darryl Brown has distinguished four general categories of nullification, three of which he views as legitimate exercises of jury decision-making power. See Brown, *supra* note 221, at 1171–96. The first occurs where the jury rejects a just law that would justly be applied to the defendant under the facts as found because government actors have transgressed. See *id.* at 1172. The second type is nullification of a law that the jury believes is generally unjust, though it would otherwise apply to the defendant under the facts as found. See *id.* at 1178–82. The third type involves a generally just law that the jury believes would be unjustly applied to the defendant in the particular case, though the facts as found would require a conviction according to the judge's instruction. See *id.* at 1183–91. Finally, the fourth type of nullification, which is the only type that Professor Brown views as outside the rule of law, is nullification based on “illegal and immoral community norms” such as racial bias. *Id.* at 1191–96. Professor Nancy Marder distinguishes cases of nullification along slightly different lines, identifying that a jury may nullify “to avoid applying a law to a particular defendant,” “to avoid applying a law that it regards as bad,” or “as a response to social conditions.” Marder, *supra* note 26, at 879.

²³² See Irwin A. Horowitz, *Jury Nullification: An Empirical Perspective*, 28 N. ILL. U. L. REV. 425, 427 (2008).

²³³ Examples found by Kalven and Zeisel in their classic study include juries' refusals to convict defendants of so-called “sumptuary” laws, including fame, liquor, gambling, and drunken driving. HARRY KALVEN, JR. & HANS ZEISEL, *THE AMERICAN JURY* 286–97 (1966). A more troubling example is the behavior of juries in rape prosecutions. See generally SUSAN ESTRICH, *REAL RAPE* (1987).

²³⁴ See Hannaford-Agor & Hans, *supra* note 45, at 1250 (“[J]ury nullification is, in essence, a counter-majoritarian measure [that] permits a small minority of citizens—twelve, or even fewer in some jurisdictions—to invalidate, in the context of a particular case, laws that have been established through the legislative process.” (footnote omitted)).

²³⁵ See, e.g., *United States v. Moylan*, 417 F.2d 1002, 1009 (4th Cir. 1969) (“To encourage individuals to make their own determinations as to which laws they will obey and which they will permit themselves as a matter of conscience to disobey is to invite chaos. No legal system could long survive if it gave every individual the option of disregarding with impunity any law which by his personal standard was judged morally untenable. Toleration of such conduct would not be democratic, as appellants claim, but inevitably anarchic.”).

²³⁶ 156 U.S. 51 (1895).

²³⁷ See, e.g., *United States v. Trujillo*, 714 F.2d 102, 106 (11th Cir. 1983) (“[D]efense counsel may not argue jury nullification during closing argument.”). Likewise, most state courts portray jury nullification as illegitimate, see, e.g., Andrew J. Parmenter, *Nullifying the Jury: “The Judicial Oligarchy” Declares War on Jury Nullification*, 46 WASHBURN L.J. 379, 388–89, 389 n.99 (2007), though two states continue to protect nullification as a right in their state constitutions, see IND. CONST. art. I, § 19; MD. CONST. (Decl. of Rights) art. XXIII (“In the trial of all criminal cases, the Jury shall be the Judges of Law, as well as of fact . . .”).

that followed over the next century, involved the issue of what a criminal jury could, or should, be told about its unquestioned power to return a verdict in opposition to the evidence and to the law as described by the trial court.²³⁸ In *Sparf*, the defendants were convicted of murder on the high seas. The trial judge rejected the defendants' proposed instruction to the effect that the jury was entitled to return a verdict on the lesser included charge of manslaughter, and instead instructed the jury that

it may be in the power of the jury, under the indictment by which these defendants are accused and tried of finding them guilty of a less crime than murder, to wit, manslaughter, or an attempt to commit murder, yet, as I have said in this case, if a felonious homicide has been committed at all, of which I repeat you are the judges, there is nothing to reduce it below the grade of murder.²³⁹

Perhaps unsurprisingly, the jury appeared confused by this instruction and returned after some initial deliberation to request clarification about whether "the defendants can be found guilty of manslaughter, or that the defendants must be found guilty" of murder.²⁴⁰ The colloquy with the trial court included the following:

Juror: Then, as I understand your honor clearly, there is nothing about manslaughter in this court?

Court: No; I do not wish to be so understood. A verdict must be based on evidence, and in a proper case a verdict for manslaughter may be rendered.

Juror: A crime committed on the high seas must have been murder, or can it be manslaughter?

Court: In a proper case, it may be murder, or it may be manslaughter but in this case it cannot be properly manslaughter. . . . Do not understand me to say that manslaughter or murder has been

²³⁸ This power stems from the general nature of a verdict in a criminal case, the double jeopardy clause of the Fifth Amendment, the inability of courts to override acquittals, and the rule of *Bushel's Case* to the effect that a juror may not be punished for returning a verdict in opposition to the instructions of the court. There is little debate about the power of the jury to nullify; the debate concerns whether the jury has a corresponding right to nullify and further whether it should be informed of such a right. *But see* Andrew D. Leipold, *Rethinking Jury Nullification*, 82 VA. L. REV. 253, 267–76 (1996) (arguing that Fifth Amendment double jeopardy protection should not preclude appellate review of all criminal acquittals, and that in general the Constitution should not be interpreted to protect the jury's power to nullify).

²³⁹ *Sparf*, 156 U.S. at 60 (emphasis omitted).

²⁴⁰ *Id.* at 61 n.1. During a dialogue with the trial judge as quoted in the opinion, a juror said, "I am the spokesman for two of us. We desire to clearly understand the matter. It is a barrier in our mind to our determining the matter. . . . [A]s I understand, it must be one thing or the other. It must be guilty or not guilty." *Id.*

committed. That is for you gentlemen to determine from the testimony and the instructions I have given you.²⁴¹

The juror then asked about the penalty for murder, and was told that it was capital punishment. He again asked the judge for clarification about whether the jury was required either to convict the defendant of capital murder or acquit altogether. The court responded:

In a proper case, a verdict for manslaughter may be rendered, . . . *and even in this case you have the physical power to do so*; but as one of the tribunals of the country, a jury is expected to be governed by law, and the law it should receive from the court.²⁴²

The juror then said, “Now it is clearly interpreted to us, and no doubt we can now agree on certain facts.”²⁴³ The jury returned a verdict of murder.

Within the jury nullification debate, *Sparf* is widely understood to have effectively sounded the death knell of the jury’s right to determine the law.²⁴⁴ The Supreme Court very explicitly demarcated the line between the function of the court and that of the jury:

We must hold firmly to the doctrine that in the courts of the United States it is the duty of juries in criminal cases to take the law from the court, and apply that law to the facts as they find them to be from the evidence. Upon the court rests the responsibility of declaring the law; upon the jury, the responsibility of applying the law so declared to the facts as they, upon their conscience, believe them to be.²⁴⁵

In the century or so since the case was decided, federal courts have taken very seriously the notion that though the jury may unfortunately have the power to disregard the law as declared by the court, it has no right to do so.²⁴⁶ Thus, where trial courts have had an opportunity to constrain the power of

²⁴¹ *Id.* at 62 n.1 (emphasis omitted).

²⁴² *Id.* (emphasis added) (emphasis omitted). As the italicized language makes clear, the jury in *Sparf* was told of its power to nullify the law; nevertheless, it was instructed that it should not exercise this power.

²⁴³ *Id.*

²⁴⁴ *See, e.g.,* Parmenter, *supra* note 237, at 388.

²⁴⁵ *Sparf*, 156 U.S. at 102. Whereas the concept of “conscience” in the nullification debate is generally assumed to imply a non-rational (or supra-rational) rejection of the facts in favor of some notion of higher justice or mercy, here the Court employed the term to suggest the jurors’ true understanding of the facts.

²⁴⁶ The Supreme Court has yet to revisit the issue of jury nullification; its lone foray has been a footnote in a case that concerned the bounds of the prosecutor’s power to introduce certain facts in a relatively prejudicial form. *See* *Old Chief v. United States*, 519 U.S. 172 (1997). For a discussion of the relevance of *Old Chief* to the jury nullification debate, see Todd E. Pettys, *Evidentiary Relevance, Morally Reasonable Verdicts, and Jury Nullification*, 86 IOWA L. REV. 467 (2001).

juries to nullify the law, they have often done so;²⁴⁷ likewise, appellate courts have upheld these judicial actions.²⁴⁸ For example, courts have excused potential jurors who had been exposed to jury nullification pamphlets, those who admitted on voir dire to belonging to organizations that espouse jury nullification, and those who admitted simply being aware of such organizations.²⁴⁹ In a sign of how far courts are willing to go to discourage jurors from engaging in nullification, the Court of Appeals for the Second Circuit, in *United States v. Thomas*, held that a deliberating juror's intention to nullify amounted to misconduct that could constitute "just cause" for dismissal of the juror prior to the verdict.²⁵⁰ Thus, where other jurors complained to the trial judge that one among them was refusing to follow the judge's instructions, it was permissible for the trial court to interview the jurors to determine whether the juror was engaged in such "a violation of [his] sworn duty."²⁵¹

Thus far, though a very few scholars have proposed reducing or eliminating the jury's *power* to nullify,²⁵² most courts and scholars have presumed the power and instead debated the contours of the right. Notwithstanding the Second Circuit's decision in *Thomas*, most instances of alleged nullification do not come to the court's attention prior to the jury's issuance of a verdict of "not guilty" in the face of strong evidence of guilt. Therefore, the modern debate over jury nullification has mainly centered upon the issue of jurors' knowledge of their nullification power, and in particular on the extent to which trial jurors may or should be informed of their power to reject the law as instructed by the court.²⁵³ The primary argument against informing juries of their power to

²⁴⁷ See Noah, *supra* note 227, at 1621.

²⁴⁸ *Id.*

²⁴⁹ The Fully Informed Jury Association (FIJA) has been active in disseminating information about the jury's nullification power. See <http://www.fija.org/> (last visited Nov. 18, 2008).

²⁵⁰ *United States v. Thomas*, 116 F.3d 606 (2d Cir. 1997). In *Thomas*, the trial court dismissed the sole black juror in a drug prosecution trial in which all of the defendants were black. Though the Second Circuit held that a juror's intent to disregard the court's instructions on the law was "just cause" for dismissal under Federal Rule of Criminal Procedure 23(b), the court further held that "the need to safeguard the secrecy of jury deliberations requires the use of a high evidentiary standard for the dismissal of a *deliberating* juror" for such a reason, and that this high standard had not been met in the case where the juror's stated reasons for his position were ambiguous and permitted an interpretation inconsistent with nullificatory intent. *Id.* at 618.

²⁵¹ *Id.* at 616.

²⁵² For one such proposal, see Leipold, *supra* note 238, at 253 (arguing that the criminal jury's unconstrained nullification power leads to legal and factual error, and proposing reforms including prosecution appeals from jury acquittals).

²⁵³ This debate came to a head following the publication of a controversial article by Professor Paul Butler, in which he advocated nullification by black jurors in cases involving nonviolent crimes committed by black defendants. See Paul Butler, *Racially Based Jury Nullification: Black Power in the Criminal Justice System*, 105 YALE L.J. 677 (1995); see also Andrew D. Leipold, *The Dangers of Race-Based Jury*

nullify is that argument or instruction about this power will lead jurors to exercise it more frequently, and in less worthy circumstances.²⁵⁴ Recent empirical research suggests that the issue is complicated, and that in some cases a nullification instruction may indeed make juries more likely to nullify.²⁵⁵

If jurors become aware that they have the power to acquit against the evidence and against the court's instructions on the law, it is possible that this knowledge could influence their verdicts. The incidence of nullification might conceivably be affected either in quantity—juries would nullify more frequently; or in type—juries would nullify in different kinds of cases than if they were not instructed on the power.²⁵⁶ Furthermore, it is unclear whether the source of this knowledge matters: Does learning of the nullification power through a pamphlet or a story on the nightly news have the same effect as being instructed by a judge, or hearing argument from a defense attorney?²⁵⁷ Arguably, being instructed by the judge that it is within their power to return a verdict contrary to the court's legal instruction is a more powerful signal than simply becoming aware of the power through another source.²⁵⁸

Nullification: A Response to Professor Butler, 44 UCLA L. REV. 108 (1996); Paul Butler, *The Evil of American Criminal Justice: A Reply*, 44 UCLA L. REV. 143 (1996).

²⁵⁴ See, e.g., *United States v. Dougherty*, 473 F.2d 1113, 1135 (D.C. Cir. 1972); Erick J. Haynie, *Populism, Free Speech, and the Rule of Law: The "Fully Informed" Jury Movement and Its Implications*, 88 J. CRIM. L. & CRIMINOLOGY 343, 361 (1997) (arguing that "the ultimate effect of nullification instructions" would be to give jurors "an open invitation to frustrate the policies of Congress or the state legislatures"); Alan Schefflin & Jon Van Dyke, *Jury Nullification: The Contours of a Controversy*, LAW & CONTEMP. PROBS., Autumn 1980, at 51, 98–102 (characterizing this argument as "The 'Damn-Good-Reason' Position").

²⁵⁵ See Norbert L. Kerr et al., *Jury Nullification Instructions as Amplifiers of Bias*, INT'L COMMENT. ON EVIDENCE, Apr. 2008, <http://www.bepress.com/cgi/viewcontent.cgi?article=1068&context=ice> (finding that nullification instructions might amplify emotional biases in cases in which the fairness of the law is in question, but that the direction and complexity of the effect requires further study); Horowitz et al., *supra* note 22 (reviewing empirical literature on nullification).

²⁵⁶ Empirical research suggests that the amount and direction of nullification is affected by the type of crime, certain features of the defendant and victim, and the type of instruction given to the jury. See Kerr et al., *supra* note 255; Horowitz et al., *supra* note 22; Horowitz, *supra* note 232 (summarizing empirical research on jury nullification).

²⁵⁷ Others have suggested that, given the likelihood that jurors will learn of their power to nullify through these non-official channels, it is preferable that courts give them accurate and complete information. See, e.g., Alan W. Schefflin & Jon M. Van Dyke, *Merciful Juries: The Resilience of Jury Nullification*, 48 WASH. & LEE L. REV. 165 (1991); cf. Stephan Landsman, *Of Mushrooms and Nullifiers: Rules of Evidence and the American Jury*, 21 ST. LOUIS U. PUB. L. REV. 65, 77 (2002) ("[S]ome number of jurors will have personal information regarding an embargoed subject . . . Unfortunately, there is no guarantee that this information is accurate.").

²⁵⁸ Like the experimenter's direction to the subject to lie in the neuroimaging studies described in Part II, *supra*, such tacit approval implied by the judge's instruction could affect the jury's decision.

Jurors presented with more and more cases in which the facts were all but certain would be placed in the logical position of believing either that their function was simply to rubber-stamp the prosecution's demand for conviction²⁵⁹ or that they were entitled and expected to exercise their non-fact-finding functions. Under these circumstances, the system would be forced to confront directly the legitimacy of the moral, merciful, and checking functions of the jury that, at present, tend to be obscured by the focus on jury fact-finding. At that point, there would be presented the genuine issue whether the jury is merely a fact-finding machine that should be replaced by a better fact-finding machine, or rather whether it has a political, institutional role worth preserving apart from its ability to judge the credibility of witnesses and the historical facts.

Recent Supreme Court cases that have addressed the role of the criminal jury complicate the analysis.²⁶⁰ The Court has focused both on the fact-finding function of the jury and on its institutional checking function. Indeed, the Court has protected the fact-finding role of the jury as a *means* of preserving its institutional checking role.²⁶¹ Historically, ensuring jury independence in fact-finding would have had the effect of preserving the jury as a bulwark against prosecutorial overreaching.²⁶² However, in a world in which facts were reliably determined by machines, the two would no longer work in tandem. At that moment, we would be faced with the constitutional choice whether it is the jury's province as fact finder that is foundational, or rather whether the more fundamental constitutional consideration is the jury's status as an independent political actor.²⁶³ The former would point in the direction of excluding even

²⁵⁹ One salutary effect of a foolproof lie detection device would be (one would hope) that fewer cases involving factually innocent defendants would be brought to trial.

²⁶⁰ See *supra* note 130; see also Arie M. Rubenstein, Note, *Verdicts of Conscience: Nullification and the Modern Jury Trial*, 106 COLUM. L. REV. 959 (2006) (arguing that modern Supreme Court cases reveal a move toward a functionalist construction of the right to trial by jury).

²⁶¹ See Seaman, *supra* note 44, at 866–67.

²⁶² For example, *Bushell's Case*, (1670) 124 Eng. Rep. 1006 (C.P.), is frequently invoked as the basis for the jury's right to decide the law. In that case, the court held that the jurors could not be punished for failing to acquit William Penn and William Mead as demanded by the court. However, as several scholars have noted, Justice Vaughan's opinion actually protects the jury not by preserving its power to determine the law, but by protecting its power to find facts. See CONRAD, *supra* note 40, at 27 (noting that "[t]he opinion was premised on the presumption that the jurors acquitted Penn and Mead because they disagreed with the court on the facts, and that the court had no authority to decide the facts of the case"); Matthew P. Harrington, *The Law-Finding Function of the American Jury*, 1999 WIS. L. REV. 377, 384 & n. 24.

²⁶³ Recently, Judge Weinstein addressed this point in the context of a defendant's right to have the court inform the jury of the sentencing consequences of its verdict. See *United States v. Polizzi*, 549 F. Supp. 2d 308 (E.D.N.Y. 2008). Based on an exhaustive analysis, Judge Weinstein concluded as follows:

reliable evidence if it threatened the jury's role as fact finder. The latter would argue for giving the jury the benefit of all reliable factual evidence and allowing it forthrightly to exercise its legislative role where it sees fit to do so.

CONCLUSION: PANDORA'S JURY BOX

Inherent in the metaphor of the black box is the notion that the process that takes place inside is ultimately unknowable, and that this very unknowability is a critical ingredient in the process. If the earliest trials were viewed as revealing the hand of divine justice, the modern view of the jury's commonsense ability to dispense justice might be seen as incorporating something akin to a religious revelation.²⁶⁴ Indeed, one way of understanding the metaphor of the black box is that it is grounded in faith that the "correct" or "true" answer will magically emerge though no reason must or even can be given.²⁶⁵ Faith is often defined as that which begins where reason ends. The hallmark of the general verdict is that it is singular, whole, and not explained nor explainable by reason. This is what distinguishes a judge's decision from a jury's decision.²⁶⁶ Professor Fisher has stated that, despite all its advantages as

[T]he emphasis on originalism by the Supreme Court in sentencing and confrontation requires enforcement of a basic element of the Sixth Amendment as originally understood: the jury of the vicinage, being aware of the sentencing implications of a finding of guilt, had the frequently exercised power to refuse to follow the law as construed by the court, and could acquit or downgrade the crime in order to avoid a sentence it deemed excessive.

Id. at 322.

²⁶⁴ See Ruprecht, *supra* note 164, at 247 n.128 ("Edson R. Sunderland first compared modern jury secrecy to ancient oracles: 'The general verdict is as inscrutable and essentially mysterious as the judgment which issued from the ancient oracle at Delphi. . . . The court protects the jury from all investigation and inquiry as fully as the temple authorities protected the priestess who spoke to the suppliant votary at the shrine.'" (quoting Edson R. Sunderland, *Verdicts, General and Special*, 29 YALE L.J. 253, 258 (1920))); Abraham S. Goldstein, *Jury Secrecy and the Media: The Problem of Postverdict Interviews*, 1993 U. ILL. L. REV. 295, 295 ("Prying into the jury's verdict, said Holdsworth, would have been as 'impious' as questioning the judgments of God.").

²⁶⁵ See, e.g., Alison Markovitz, Note, *Jury Secrecy During Deliberations*, 110 YALE L.J. 1493, 1505 (2001) ("Early notions that the jury should deliberate in secret were linked to the conception of the jury as an enigmatic, divinely inspired body. . . . This religious or mystical origin still resonates in current discussions of jury secrecy."). It is interesting that some early judicial discussions of scientific and statistical evidence—exactly the sorts of evidence that would tend to diminish the need for jury fact-finding—employed rhetorical imagery of sorcery and magic. See *People v. Collins*, 438 P.2d 33, 33 (Cal. 1968) (calling mathematics "a veritable sorcerer").

²⁶⁶ Cf. Rachel E. Barkow, *The Ascent of the Administrative State and the Demise of Mercy*, 121 HARV. L. REV. 1332, 1334 (2008) ("[T]he rise of the administrative state has made unchecked discretion an anomaly in the law, and a phenomenon to be viewed with suspicion."). Some scholars have proposed reforms that would require juries to give reasons for their verdicts under certain circumstances. See John D. Jackson, *Making*

a source of legitimacy, “[o]ne source of legitimacy the jury could never comfortably claim is divinity.”²⁶⁷ And yet it does seem that one attraction of a black box—whether the jury or the mind—is its inscrutability and the sense that it cannot be explained purely by logic nor broken down into its component parts.²⁶⁸

Yet just as the black box of the mind has begun to admit some light through the use of developing brain-imaging technologies, so the black box of the jury room may ultimately relinquish a bit of its inscrutability. This Article has suggested that one consequence of advances in scientific techniques is that historical fact in many cases will become more certain. In particular, were an accurate lie detector developed, the jury’s unique role in determining witness credibility would be called into question. At that moment, in many cases, the criminal jury would exist either to rubber-stamp the prosecution’s version of the historical facts, or instead to serve distinct, non-fact-finding functions in addition to determining historical fact. The former vision of the jury would relegate it to an increasingly trivial role, one at odds with the history, precedent, and purposes of the right to jury trial embodied in the Constitution. The latter vision, in contrast, would allow the jury knowledge of its legislative function, even while risking that it might exercise it in ways that courts or the public find illegitimate. Perhaps the most that can be said at present is this: the choice between alternate visions of the jury does not cease to exist simply because it is currently hidden.

Juries Accountable, 50 AM. J. COMP. L. 477 (2002) (proposing that juries be accountable for convictions, but not for acquittals); Leipold, *supra* note 238, at 321 (suggesting that special verdicts be employed in criminal cases in which “a judge thought that the jury would be tempted to stray beyond the boundaries of the law”).

²⁶⁷ Fisher, *supra* note 30, at 705.

²⁶⁸ One source of resistance to brain science is the idea that the “mind” and the “brain” are distinct; that the mind is more than simply the sum of the physical, chemical, and electrical processes that can be measured by sophisticated machinery—there is a “ghost in the machine.” See STEVEN PINKER, *THE BLANK SLATE: THE MODERN DENIAL OF HUMAN NATURE* 8–11 (2002) (describing the mind/brain controversy).