

ESSAY

CAN BAD SCIENCE BE GOOD EVIDENCE? NEUROSCIENCE, LIE DETECTION, AND BEYOND

Frederick Schauer†

INTRODUCTION

How should the legal system confront the advances in the brain sciences that may possibly allow more accurate determinations of veracity—lie detecting—than those that now pervade the litigation process? In this Essay, I question the view, widespread among the scientists most familiar with these advances, that the neuroscience of lie detection is not, or at least not yet, nearly reliable enough to be used in civil or criminal litigation or for related forensic purposes. But in challenging the neuroscientists and their allies, I make no claims about the science of lie detection that go beyond the current state of scientific knowledge or, more importantly, my own ability to speak about the relevant scientific developments. Rather, I argue that because law's goals and norms differ from those of science, there is no more reason to impose the standards of science on law than to impose the standards of law on science. Law must use science, and should always prefer good science to bad. In some contexts, however, good science may still not be good enough for law, while in other contexts—hence the title of this Essay—bad science, as measured by the standards of scientists, may still have valuable legal uses. To be clear, my goal in this Essay is decidedly *not* to argue that neuroscience-based lie detection *should*, now or even in the foreseeable future, necessarily be admissible in court or used for other forensic purposes. Rather,

† David and Mary Harrison Distinguished Professor of Law, University of Virginia. This Essay was presented at the Mini-Foro on Proof and Truth in the Law, Institute for Philosophical Research, Universidad Nacional Autonoma de México (UNAM), as an Inaugural Lecture at the University of Virginia School of Law, at the Duck Conference on Social Cognition, to the Departments of Psychology at Michigan State University and the University of Virginia, and at the Baldy Center for Law and Social Policy at the University of Buffalo School of Law. Many of the ideas presented here were generated during meetings of the John D. and Catherine T. MacArthur Foundation's Law and Neuroscience Project, whose tangible and intangible support I gratefully acknowledge. Detailed and constructive comments by Charles Barzun, Teneille Brown, Jonathan Gradowski, Joshua Greene, Owen Jones, Adam Kolber, Greg Mitchell, John Monahan, Peter Tillers, Lloyd Snook, and Bobbie Spellman have made this version far better than its predecessors.

my goal is to argue that the question of whether the law should use neuroscience-based lie detection cannot be answered by scientific standards of reliability and validity alone. Science can—and should—inform the legal system about facts, including facts about degrees of reliability and the extent of experimental validity, but the ultimate normative and institutional question of whether and when, if at all, a given degree of validity or reliability is sufficient for some legal or forensic purpose is a legal and not a scientific question.

In important respects, this analysis of the potential legal uses of neuroscience-based lie detection is more a case study than a discrete topic. Most of what I argue here applies to other forms of lie detection, to other forms of scientific evidence, and indeed to evidence generally. Thus, as I elaborate in Part V of this Essay, my central theme calls into doubt important dimensions of the modern revolution in the standards for the admission of scientific evidence. Commencing with *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,¹ and continuing through *General Electric Co. v. Joiner*² and *Kumho Tire Co., Ltd. v. Carmichael*,³ the Supreme Court has for almost two decades attempted to deal with the very real problem of “junk science” by imposing increasingly stringent scientific standards of reliability and experimental validity on the admissibility of scientific evidence and expert testimony in the federal courts.⁴ By dealing with science and experts but not with the myths and superstitions that pervade the fact-finding process, however, the Court may have unintentionally lowered the quality of evidence generally. By discouraging poor science while leaving nonscience untouched, the *Daubert* revolution may have perversely fostered an increased reliance on the even worse nonscientific evidence that dominates the litigation process by not masquerading as science at all. This problem may not have an easy solution, but its identification suggests that *Daubert* may have created as many problems as it solved. The revolution in scientific and expert testimony that started with *Daubert* is thus—or at least should be—far from over.

¹ 509 U.S. 579 (1993).

² 522 U.S. 136 (1997).

³ 526 U.S. 137 (1999).

⁴ See *infra* note 24 and accompanying text. Numerous state courts have also adopted the *Daubert* approach. See, e.g., *State v. Porter*, 698 A.2d 739, 742 (Conn. 1997); *Commonwealth v. Lanigan*, 641 N.E.2d 1342, 1348–49 (Mass. 1994); *Miller v. Bike Athletic Co.*, 687 N.E.2d 735, 740–41 (Ohio 1998); *Wilt v. Buracker*, 443 S.E.2d 196, 203 (W. Va. 1993).

I

LIE DETECTION, THEN AND NOW: THE CLAIMS

I begin by describing the current controversy over the legal and forensic⁵ uses of neuroscience-based lie detection. In some respects, this controversy should come as no surprise. The common law litigation process places huge reliance on the sworn testimony of witnesses, a phenomenon that is itself worthy of note.⁶ After all, many other methods of factual investigation employ dramatically different approaches that rely far more heavily on primary, rather than secondary, sources of knowledge. The scientist who seeks to determine whether drinking red wine reduces the likelihood of heart disease does not, for example, summon representatives of the wine industry and the Temperance League to each make their cases and thereafter decide which of the two advocates is more believable. Rather, she engages in the kind of primary research that we call experimentation. Similarly, historians who conduct archival research, psychologists who experiment on subjects, empirical economists who perform multiple regressions with large data sets, oceanographers who explore the sea with scientific instruments or submersible watercraft, and researchers for policymakers who combine various techniques to determine the factual terrain that a policy will affect all engage in one form or another of primary research.⁷

Once we grasp the diverse array of primary techniques for determining facts, we can understand how unusual the legal system is not only in routinely using party-generated witnesses to provide information as to which they, but not the trier of fact, have first-hand knowledge,⁸ but also in precluding the trier of fact from obtaining the first-

⁵ In this Essay I use *forensic*, as distinguished from *legal*, to refer to those dimensions of criminal investigation that either precede a trial or exist apart from it.

⁶ Although sworn witness testimony also plays a large role in the civil law, the qualification to the common law in the accompanying text is a function of the somewhat larger role that judges in many civil law countries play in managing the process of direct factual investigation, especially in criminal cases. For good descriptions of the civil law approach, see EUROPEAN CRIMINAL PROCEDURES (Mireille Delmas-Marty & J.R. Spencer eds., 2002); JACQUELINE HODGSON, FRENCH CRIMINAL JUSTICE: A COMPARATIVE ACCOUNT OF THE INVESTIGATION AND PROSECUTION OF CRIME IN FRANCE (2005).

⁷ To be sure, law is not entirely alone in its reliance on testimonial evidence. Testimony and other forms of indirect evidence do play a significant role in other investigative and fact-finding enterprises. See, e.g., C.A.J. COADY, TESTIMONY: A PHILOSOPHICAL STUDY 233–76 (1992) (surveying the use of testimony in history, mathematics, and psychology); John Hardwig, *The Role of Trust in Knowledge*, 88 J. PHIL. 693 (1991) (discussing the role of testimonial evidence in science and mathematics). Nevertheless, law is noteworthy in relying on testimony and authority more than most other disciplines, and conversely, relying substantially less on direct investigation and experimentation, especially in courtroom settings. See RICHARD A. POSNER, THE PROBLEMS OF JURISPRUDENCE 62 (1990) (“[A]uthority and hierarchy play a role in law that would be inimical to scientific inquiry.”).

⁸ FED. R. EVID. 602.

hand knowledge that in other domains remains the gold standard for empirical reliability.⁹ Still, the legal system we have, idiosyncratic as it is within the realm of empirical inquiry in relying so heavily on second-hand knowledge, is one that often makes it important to determine which of two opposing witnesses is telling the truth.

Of course, not all trials involve a conflict between a truth teller and a liar. Honest misperceptions and more or less honest omissions, exaggerations, shadings, fudgings, slantings, bendings, and hedging¹⁰ are an omnipresent feature of modern litigation. But so too is flat-out lying. Because the legal system relies far more heavily on the reports of witnesses than on primary investigation by the trier of fact, it should come as no surprise that the law is preoccupied with trying to assess whether the witnesses who testify in court (or otherwise provide information for legal or forensic decision making) are telling the truth.¹¹

Historically, the law relied on the oath to serve the truth-warranting function.¹² When people genuinely believed that lying under oath would send them to hell, the law could comfortably rely on a witness's fear of eternal damnation to provide confidence that witnesses were likely to tell the truth.¹³

⁹ “‘Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry.’” *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 593 (1993) (quoting Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substance Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 Nw. U. L. Rev. 643, 645 (1992)). See also SHEILA JASANOFF, SCIENCE AT THE BAR: LAW, SCIENCE, AND TECHNOLOGY IN AMERICA 7–11 (1995) (comparing the methods and goals of law and science); Susan Haack, *Irreconcilable Differences? The Troubled Marriage of Science and Law*, 72 LAW & CONTEMP. PROBS. 1, 2, 7–15 (2009) (distinguishing “the investigative character of science” from the “adversarial culture of our legal system”); Harry W. Jones, *Legal Inquiry and the Methods of Science, in LAW AND THE SOCIAL ROLE OF SCIENCE* 122, 124 (Harry W. Jones ed., 1967) (“Legal propositions have their origin not in empirical observation but in authoritative pronouncement by a court or legislature.”); James A. Martin, *The Proposed ‘Science Court,’* 75 MICH. L. REV. 1058, 1062–64 (1977) (comparing the scientific and legal methods of determining truth).

¹⁰ See Frederick Schauer & Richard Zeckhauser, *Paltering, in DECEPTION: FROM ANCIENT EMPIRES TO INTERNET DATING* 38, 38–41 (Brooke Harrington ed., 2009) (analyzing various forms of deception not involving explicit falsity).

¹¹ See, e.g., Elizabeth B. Ford, *Lie Detection: Historical, Neuropsychiatric and Legal Dimensions*, 29 INT'L J. L. & PSYCHIATRY 159, 165 (2006) (“The detection of liars and lies has been an element of many professions in the legal [and] criminal justice . . . fields.”).

¹² See, e.g., Haack, *supra* note 9, at 3 (describing law’s early reliance on the oath and related religious tests).

¹³ See George Fisher, *The Jury’s Rise as Lie Detector*, 107 YALE L.J. 575, 580, 583 (1997) (noting that the legal system’s reliance on the oath was based on “the perceived divine power of the oath to compel truthful testimony” through “the threat of divine vengeance” for perjured testimony). Thomas Raeburn White and Daniel Blau describe the history of the oath and its religious aspects. See generally Thomas Raeburn White, *Oaths in Judicial Proceedings and Their Effect upon the Competency of Witnesses*, 51 AM. L. REG. 373 (1903); Daniel Blau, Note, *Holy Scriptures and Unholy Strictures: Why the Enforcement of a Religious Orthodoxy in*

As religious belief diminished, or at least as law's confidence in it as a guarantor of truth waned, the legal system increasingly relied on faith in the lie-exposing powers of vigorous cross-examination. As celebrated in the *Perry Mason* television series of the 1950s and 1960s and since reinforced by numerous items of popular culture,¹⁴ the legal system has long believed that cross-examination so reduces the effectiveness of lying so much that a truth-determining system that relies on witness testimony and cross-examination will not be unacceptably vulnerable to intentional deception.¹⁵

More importantly, and because cross-examination is far less effective in exposing lies and liars than television writers and viewers believe,¹⁶ the legal system has placed its faith in judges and juries. Now, the legal system assigns the task of determining veracity, and credibility in general, to the trier of fact: most visibly, even if not the most frequently, the jury.¹⁷ Among other things, the jury must assess the demeanor of witnesses, their past record of truth telling, the internal coherence of their stories, and the external coherence of their stories with the stories of others, all in order to determine who is telling the truth and who is not.¹⁸

North Carolina Demands a More Refined Establishment Clause Analysis of Courtroom Oaths, 4 FIRST AMENDMENT L. REV. 223, 227–29 (2006).

¹⁴ See Naomi Mezey & Mark C. Niles, *Screening the Law: Ideology and Law in American Popular Culture*, 28 COLUM. J.L. & ARTS 91, 116–17 (2005) (“[P]erry Mason] had such a strong popular cultural influence that millions of people born after its last episode aired are still familiar both with the character and with his signature talent for dramatic and successful cross-examination.”).

¹⁵ See 5 WIGMORE, EVIDENCE § 1367 (Chadbourn rev. 1974) (asserting that cross-examination is the “greatest legal engine ever invented for the discovery of truth”).

¹⁶ See Jules Epstein, *The Great Engine That Couldn’t: Science, Mistaken Identifications, and the Limits of Cross-Examination*, 36 STETSON L. REV. 727, 774–82 (2007) (noting the inability of cross-examination to “undercut eyewitness reliability” in some respects, which may in turn lead to erroneous identifications); Marvin E. Frankel, *The Search for Truth: An Umpireal View*, 123 U. PA. L. REV. 1031, 1036 (1975) (characterizing the truth-determining ability of the adversarial process as “untested” and lawyers’ belief in it as “self-congratulat[ory]”).

¹⁷ See, e.g., United States v. Scheffer, 523 U.S. 303, 313 (1998) (“[T]he jury is the lie-detector.”) (quoting United States v. Barnard, 490 F.2d 907, 912 (9th Cir. 1973)); United States v. Thompson, 615 F.2d 329, 332 (5th Cir. 1980) (describing the jury’s function as “credibility determination”); Bloom v. People, 185 P.3d 797, 807 (Colo. 2008) (“[P]olygraph evidence [is inadmissible because it] will prejudice the jury’s evaluation of a witness’s credibility.”); State v. Christiansen, 163 P.3d 1175, 1180 (Idaho 2007) (“[J]urors are the judges of the credibility of witnesses.”) (quoting People v. Barnes, 9 P. 532, 533 (Idaho 1886)); State v. Myers, 382 N.W.2d 91, 95 (Iowa 1986) (“[W]eighing the truthfulness of a witness is a matter reserved exclusively to the fact finder.”); State v. Lyon, 744 P.2d 231, 240 (Or. 1987) (Linde, J., concurring) (“The cherished courtroom drama of confrontation, oral testimony and cross-examination is designed to let a jury pass judgment on [parties’ and witnesses’] truthfulness and on the accuracy of their testimony.”). See generally Fisher, *supra* note 13 (offering a comprehensive historical account of the rise of the jury as lie detector in the American legal system).

¹⁸ See, e.g., James P. Timony, *Demeanor Credibility*, 49 CATH. U. L. REV. 903, 907–13 (2000) (discussing the various criteria that constitute credibility evidence, including de-

Enter science. Because the criteria that judges and juries traditionally employ to evaluate the veracity of witnesses have been notoriously unreliable,¹⁹ the quest for a scientific way of distinguishing the truth teller from the liar has been with us for generations. Indeed, the *Frye* test,²⁰ which for many years was the prevailing legal standard for determining the admissibility of scientific evidence, arose in 1923 in the context of an unsuccessful attempt to admit into evidence a rudimentary lie-detection machine invented by William Moulton Marston²¹—perhaps better known as the creator of the comic book character Wonder Woman, whose attributes included possession of a magic lasso, forged from the Magic Girdle of Aphrodite, which would make anyone it encircled tell the truth without fail.²² The device at issue in *Frye* was a simple polygraph and not a magic lasso, but *Frye* did not just set the standard for the admission of scientific evidence for more than a half-century; its exclusion of lie-detection technology also paved the way for the continuing exclusion, with few exceptions, of lie-detection evidence in American courts.²³

The science of lie detection has improved considerably since 1923, but not by so much as to have led to large-scale changes in judicial attitudes. Indeed, even after *Daubert v. Merrell Dow Pharmaceuticals*,

meanor, reputation for truthfulness, prior inconsistent statement, contradiction, and corroboration).

19 See generally ALDERT VRIJ, DETECTING LIES AND DECEIT: THE PSYCHOLOGY OF LYING AND THE IMPLICATIONS FOR PROFESSIONAL PRACTICE (2000); Jeremy A. Blumenthal, *A Wife of the Hands, a Lick of the Lips: The Validity of Demeanor Evidence in Assessing Witness Credibility*, 72 NEBR. L. REV. 1157, 1190–97 (1993); Olin Guy Wellborn III, *Demeanor*, 76 CORNELL L. REV. 1075, 1082–88 (1991).

20 *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923) (holding that scientific evidence must use methods generally accepted in the relevant scientific community in order to be admissible).

21 Don Grubin & Lars Mardin, *Lie Detection and the Polygraph: A Historical Review*, 16 J. FORENSIC PSYCHIATRY & PSYCHOL. 357, 359–60 (2005).

22 NATIONAL RESEARCH COUNCIL, THE POLYGRAPH AND LIE DETECTION 295 (2003).

23 Polygraph evidence is per se inadmissible in many jurisdictions. See, e.g., *United States v. Gilliard*, 133 F.3d 809, 815 (11th Cir. 1998); *United States v. Sanchez*, 118 F.3d 192, 197 (4th Cir. 1997); *State v. Jones*, 753 N.W.2d 677, 690 (Minn. 2008). Other jurisdictions condition the admission of polygraph evidence on the stipulation of the parties. See, e.g., CAL. EVID. CODE § 351.1 (1995); *Brown v. Darcy*, 783 F.2d 1389, 1395 (9th Cir. 1986); *Wilkins v. State*, 190 P.3d 957, 970 (Kan. 2008). New Mexico is the most prominent exception. See N.M. R. EVID. 11-707 (permitting the admission of polygraph evidence to prove the “the truthfulness of any person called as a witness”). But other jurisdictions also express some increased sympathy to polygraph use. See, e.g., *United States v. Cordoba*, 104 F.3d 225, 228 (9th Cir. 1997) (holding that a per se exclusion of polygraph evidence does not survive *Daubert*); *Rupe v. Wood*, 93 F.3d 1434, 1437 (9th Cir. 1996) (allowing defendant to use polygraph evidence at sentencing hearing); *United States v. Galbreth*, 908 F. Supp. 877, 896 (D.N.M. 1995) (allowing polygraph evidence if examiner is properly qualified); *Commonwealth v. Duguay*, 720 N.E.2d 458, 463 (Mass. 1999) (requiring the reliability of polygraphic evidence to “be established by proof in a given case that a qualified tester who conducted the test had in similar circumstances demonstrated . . . the high level of accuracy of the conclusions that the tester reached in those tests”).

*Inc.*²⁴ replaced *Frye*'s "general acceptance" test in federal courts by insisting on various indicia of scientific validity as a precondition to the admissibility of evidence purporting to be scientific,²⁵ the situation with respect to lie detection has remained much the same.²⁶

What makes the foregoing important is the rapidly changing state of cognitive neuroscience—the study of human thinking using various methods of (indirectly) measuring brain activity.²⁷ The tools of modern neuroscience are numerous, but the most prominent of them is fMRI—functional magnetic resonance imaging.²⁸ Commonly called brain scanning, fMRI examination holds out the possibility of being able to determine which parts of the brain perform which cognitive tasks. Although novices seeing images of an fMRI scan sometimes believe that certain parts of the brain are "lighting up" with electrical activity when engaged in certain tasks,²⁹ what actually occurs is that the activated portion of the brain recruits more oxygenated blood cells to help it in its task. What appears to be a lit up part of the brain is actually a part that has more oxygenated hemoglobin in it than it had when it was less, or differently, cognitively engaged.³⁰

Hardly surprisingly, the development of fMRI technology has led some researchers to conclude that this technology can be effective in distinguishing liars from truth tellers.³¹ If—and it is a huge *if*—differ-

²⁴ 509 U.S. 579, 589–90 (1993) (holding that Federal Rule of Evidence 702 displaced the *Frye* standard and instructing trial judges to "screen[] . . . scientific testimony or evidence" for both legal relevance and scientific validity). The Supreme Court added further refinements in *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 147 (1999) (holding that trial courts' gatekeeping obligations under *Daubert* extend to nonscientific expert testimony), and *General Electric Co. v. Joiner*, 522 U.S. 136, 138–39 (1997) (declaring "abuse of discretion" the appropriate standard in "reviewing a trial court's decision to admit or exclude expert testimony under [Daubert]" (quoting *Commonwealth v. Stewart*, 663 N.E.2d 255 (Mass. 1996))).

²⁵ *Daubert*, 509 U.S. at 593–94.

²⁶ See CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, EVIDENCE 701–02 (4th ed. 2009) (describing continuing exclusion of most polygraph evidence). See also *supra* note 23.

²⁷ Marcus E. Raichle, *A Brief History of Human Brain Mapping*, 32 TRENDS IN NEUROSCIENCE 118, 118 (2008) ("Cognitive neuroscience combines the experimental strategies of cognitive psychology with various techniques to actually examine how brain function supports mental activities.").

²⁸ For accessible explanations of fMRI, see SCOTT A. HUETTEL, ALLEN W. SONG, & GREGORY McCARTHY, FUNCTIONAL MAGNETIC RESONANCE IMAGING (2004); Raichle, *supra* note 27, at 118–25.

²⁹ See Teneille Brown & Emily Murphy, *Through a Scanner Darkly: Functional Neuroimaging as Evidence of a Criminal Defendant's Past Mental States*, 62 STAN. L. REV. 1119, 1190 (2010); Matthew Baptiste Holloway, *One Image, One Thousand Incriminating Words: Images of Brain Activity and the Privilege Against Self-Incrimination*, 27 TEMP. J. SCI. TECH. & ENVTL. L. 141, 146 (2008).

³⁰ Holloway, *supra* note 29, at 145–46.

³¹ The literature is large and growing. Some studies claim to have produced positive conclusions regarding the capabilities of neuroscience-based lie detection. See, e.g., Christos Davatzikos et al., *Classifying Spatial Patterns of Brain Activity with Machine Learning Meth-*

ent parts of the brain are active when a person is lying than when telling the truth, or when acting deceptively rather than honestly, then brain scans might be able to determine whether a person is lying or telling the truth. Or so it is claimed. And especially by those who see the commercial potential for just this technology. For-profit com-

ods: *Application to Lie Detection*, 28 NEUROIMAGE 663, 667–68 (2005) (concluding that a nonlinear pattern classification method can detect patterns of brain activity associated with lying); G. Ganis et al., *Neural Correlates of Different Types of Deception: An fMRI Investigation*, 13 CEREBRAL CORTEX 830, 832–38 (2003) (yielding results that “show that different patterns of brain activation arise when people tell lies than when they tell the truth”); Joshua D. Greene & Joseph M. Paxton, *Patterns of Neural Activity Associated with Honest and Dishonest Moral Decisions*, 106 PROC. NAT'L ACAD. SCI. U.S. 12506, 12509–10 (2009) (suggesting that individual differences in brain “control network activity” are associated with differences in presence of dishonest behavior); F. Andrew Kozel, Tamara M. Padgett & Mark George, Brief Communication, *A Replication Study of the Neural Correlates of Deception*, 118 BEHAV. NEUROSCIENCE 852, 855 (2004) (finding that “[f]or lying, compared with telling the truth, there is more activation in the right anterior cingulate, right inferior frontal, right orbitofrontal, right middle frontal, and left middle temporal areas”); F. Andrew Kozel et al., *Detecting Deception Using Functional Magnetic Resonance Imaging*, 58 BIOLOGICAL PSYCHIATRY 605, 611 (2005) [hereinafter *Detecting Deception*] (concluding that “fMRI can be used to detect deception within a cooperative individual”); Andrew Kozel et al., *A Pilot Study of Functional Magnetic Resonance Imaging Brain Correlates of Deception in Healthy Young Men*, 16 J. NEUROPSYCHIATRY & CLINICAL NEUROSCIENCE 295, 304 (2004) (concluding that using blood oxygen level dependent fMRI “to investigate brain changes associated with deception is . . . possible”); Daniel D. Langleben et al., *Telling Truth from Lie in Individual Subjects with Fast Event-Related fMRI*, 26 HUM. BRAIN MAPPING 262, 271 (2005) [hereinafter *Telling Truth*] (concluding that fMRI images may be able to distinguish a truth from a lie on the basis that a lie “appears to be a more working memory-intensive activity, characterized by increased activation of the inferolateral cortex implicated in response selection, inhibition, and generation”); D. D. Langleben et al., Rapid Communication, *Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study*, 15 NEUROIMAGE 727, 730–31 (2002) (finding a “neurophysiological difference between deception and truth”); Tatia M.C. Lee et al., *Neural Correlates of Feigned Memory Impairment*, 28 NEUROIMAGE 305, 310–12 (2005); Tatia M.C. Lee et al., *Lie Detection by Functional Magnetic Resonance Imaging*, 15 HUM. BRAIN MAPPING 157, 161–63 (2002) (concluding that it is “unfeasible” to control one’s cerebral activity to avoid lie detection); Donald H. Marks, Mehdi Adineh & Sudeepa Gupta, *Determination of Truth From Deception Using Functional MRI and Cognitive Engrams*, 5 INTERNET J. RADIOLOGY 1 (2006), http://www.ispub.com/journal/the_internet_journal_of_radiology/volume_5_number_1_35/article/determination_of_truth_from_deception_using_functional_mri_and_cognitive_engrams.html (showing that “specific activation patterns occur in the brain of individuals looking at specific pictures, and also whether they are contemplating giving a truthful or a deceptive response”); 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, 679 (2006) (concluding that “[s]pecific areas of the brain involved in deception or truth telling can be depicted with functional MR imaging”); Jennifer Maria Nuñez et al., *Intentional False Responding Shares Neural Substrates with Response Conflict and Cognitive Control*, 25 NEUROIMAGE 267, 273–76 (2005) (finding certain brain regions to be “significantly more active when falsifying information as compared to when answering truthfully”); Sean A. Spence et al., *Speaking of Secrets and Lies: The Contribution of Ventrolateral Prefrontal Cortex to Vocal Deception*, 40 NEUROIMAGE 1411, 1415–18 (2008); Sean A. Spence et al., *Behavioural and Functional Anatomical Correlates of Deception in Humans*, 12 NEUROREPORT 2849, 2851–52 (2001) (finding that individuals telling lies have increased response times and increased activation in specific regions of the brain).

panies, in particular No Lie MRI³² and Cephos,³³ have already begun marketing their lie-detection services, and these companies and their principals have been at the forefront of those touting the courtroom and forensic potential of the new technology.

Neuroscience-based lie detection follows a long history of lie-detection technology. The earliest polygraphs analyzed blood pressure, but modern techniques include electroencephalography, which measures brain-generated electrical current,³⁴ facial microexpression analysis, developed by the psychologist Paul Ekman³⁵ and featured in the television series "Lie to Me,"³⁶ periorbital thermography,³⁷ which measures the temperature around the eyes; and near-infrared spectroscopy,³⁸ which uses infrared light to measure changes in blood flow and is thus the precursor of fMRI technology. These technologies and methods have their adherents, but I focus on fMRI because it is potentially the most reliable of these techniques—although not reliable enough, as we will see, to persuade those most familiar with the technology to endorse it for courtroom or forensic use.

II

NEUROSCIENCE-BASED LIE DETECTION: THE COUNTERCLAIMS

In legal and policy debates, perhaps almost as much as in physics, every action appears to produce an equal and opposite reaction. And so it has been with the reaction of mainstream academic neuroscientists to the claims about the lie-detection potential of fMRI scans. A prominent article by Stanford law professor Henry Greely and neuroscientist Judy Illes surveyed all of the existing studies of neuroscience-based lie detection through 2006 and concluded that the studies individually fell far short of existing scientific standards of rigor and collectively did not come close to establishing the reliability of fMRI-based lie detection.³⁹ Accordingly, Greely and Illes urged a legally

³² See No Lie MRI Homepage, <http://www.noliemri.com> (last visited Apr. 10, 2010).

³³ See Cephos Homepage, <http://www.cephoscorp.com> (last visited Apr. 10, 2010).

³⁴ See, e.g., Lawrence A. Farwell & Sharon S. Smith, *Using Brain MERMER to Detect Knowledge Despite Efforts to Conceal*, 46 J. FORENSIC SCI. 135–37, 142 (2001).

³⁵ PAUL EKMAN, TELLING LIES: CLUES TO DECEIT IN THE MARKETPLACE, POLITICS, AND MARRIAGE 123–61 (rev. ed. 2009).

³⁶ See *Lie to Me* Homepage, <http://www.fox.com/lietome/about> (last visited Apr. 10, 2010).

³⁷ See, e.g., I. Pavlidis & J. Levine, *Monitoring of Periorbital Blood Flow Rate Through Thermal Image Analysis and Its Application to Polygraph Testing*, 3 PROC. 23RD ANN. INT'L CONF. IEEE ENGINEERING MED. & BIOLOGY SOC'Y 2826, 2829 (2001).

³⁸ See, e.g., Britton Chance et al., *A Novel Method for Fast Imaging of Brain Function, Non-Invasively, with Light*, 2 OPTICS EXPRESS 411, 413 (1998).

³⁹ Henry T. Greely & Judy Illes, *Neuroscience-Based Lie Detection: The Urgent Need for Regulation*, 33 AM. J. L. & MED. 377, 395–404 (2007). See also Jane Campbell Moriarty, *Visions of Deception: Neuroimages and the Search for Truth*, 42 AKRON L. REV. 739, 758–61 (2009) (arguing that fMRI-based lie detection satisfies neither *Frye* nor *Daubert* standards); Michael S.

imposed moratorium on the use of the technology for courtroom or forensic purposes until a federal regulatory agency could establish its reliability according to scientific standards.⁴⁰ Similarly, one leading neuroscientist has insisted that the “data offer no compelling evidence that fMRI will work for lie detection in the real world.”⁴¹ Another has concluded that “[a]t present we have no good ways of detecting deception despite our very great need for them.”⁴² And still another has concluded that using laboratory findings on fMRI lie detection in settings “that can potentially impact individuals’ legal rights” should, on the current state of knowledge, “remain a research topic, instead of a legal tool.”⁴³ An editorial in *Nature Neuroscience* joined the chorus of skepticism,⁴⁴ as did a report from a National Research Council committee,⁴⁵ and several published articles by researchers and practitioners from various disciplines insisted that fMRI lie detection was not ready for the “real world.”⁴⁶

Pardo, *Neuroscience Evidence, Legal Culture, and Criminal Procedure*, 33 AM. J. CRIM. L. 301, 311–20 (2006) (detailing conceptual, empirical, and practical limitations of fMRI-based lie detection for use in the courtroom).

⁴⁰ Greely & Illes, *supra* note 39, at 413.

⁴¹ Nancy Kanwisher, *The Use of fMRI Lie Detection: What Has Been Shown and What Has Not*, in *USING IMAGING TO IDENTIFY DECEIT: SCIENTIFIC AND ETHICAL QUESTIONS* 7, 12 (2009). Kanwisher’s use of “compelling” to set her threshold for usability is noteworthy because much of the question of fMRI’s legal use depends precisely on how strong the case must be for fMRI-based lie detection before its use is permissible. To assert that the case must be compelling is thus to impose a very high burden, but why the burden should be as high as requiring “compelling” evidence, as opposed to, say, “plausible” evidence, “some” evidence, or “more persuasive than not” evidence, is unclear. Much that follows in this Essay is exactly about this issue, but it is important to recognize that those who insist at the outset that the evidence be compelling have stacked the deck by, in essence, assuming the conclusion.

⁴² Marcus E. Raichle, *An Introduction to Functional Brain Imaging in the Context of Lie Detection*, in *USING IMAGING TO IDENTIFY DECEIT*, *supra* note 41, at 3, 6.

⁴³ Elizabeth A. Phelps, *Lying Outside the Laboratory: The Impact of Imagery and Emotion on the Neural Circuitry of Lie Detection*, in *USING IMAGING TO IDENTIFY DECEIT*, *supra* note 41, at 14, 20.

⁴⁴ Editorial, *Deceiving the Law*, 11 NATURE NEUROSCIENCE 1231, 1231 (2008) (“There is little evidence to indicate that the newer [fMRI-based] lie-detection technologies . . . work well enough to detect deception accurately on an individual level with an error rate that is low enough to be anywhere near acceptable in court.”).

⁴⁵ NATIONAL RESEARCH COUNCIL, *EMERGING COGNITIVE NEUROSCIENCE AND RELATED TECHNOLOGIES* 37 (2008) (concluding that “there has not yet been sufficient systematic research to determine if functional neuroimaging can meet the challenges to the neurophysiological detection of psychological states relevant to deception”).

⁴⁶ See, e.g., James R. Merikangas, *Commentary: Functional MRI Lie Detection*, 36 J. AM. ACAD. PSYCHIATRY & L. 499, 499–501 (2008) (concluding that fMRI lie detection technology does not meet the *Daubert* criteria for courtroom testimony); Michael S. Gazzaniga, *The Law and Neuroscience*, 60 NEURON 412, 413, 415 (2008) (cautioning against the introduction of neuroscience-based lie detection evidence because jurors and judges may erroneously accept such evidence as legally dispositive); Jed S. Rakoff, *Lie Detection in the Courts: The Vain Search for the Magic Bullet*, in *USING IMAGING TO IDENTIFY DECEIT*, *supra* note 41, at 40, 44 (arguing that neuroscience-based lie detection “suffers from several defects that would render such evidence inadmissible under [Federal Rule of Evidence] 702”); Joseph R.

At the core of the campaign against the use of fMRI in real-world legal settings is the conviction that the existing state of the research is “poor science.” And it is poor science, it is said, not only because of doubts about the reliability rates of fMRI lie detection, but also, and more fundamentally, because the research that allegedly supports the proposed techniques and determines their reliability rates has serious problems of experimental validity.⁴⁷ The critics claim that the tests that have been conducted are different in material ways from real-world lying and truth telling, thus undermining the inference that fMRI detection could accurately detect real-world liars just because it accurately detects liars in experimental settings.⁴⁸ Part of the difference, indeed the major difference, is that in most instances the researchers have instructed the experimental subjects to lie. But whether an instructed lie is even a lie at all presents substantial questions of construct validity—whether the experiment measures what it purports to measure—that cast significant doubt on the research conclusions.⁴⁹ Additional doubts stem from the size and nature of the samples, potential confounding variables (e.g., whether subjects are left- or right-handed),⁵⁰ and the significant possibility that subjects could take countermeasures to render the test results unreliable.⁵¹ Thus, the existing research stands accused of being flawed even as pure laboratory research and of being far less applicable to nonlaboratory settings than its proponents have typically claimed.

The charges against the existing research go even further. The results have often been neither published in peer-reviewed journals nor replicated, thus failing to satisfy the normal standards for assess-

Simpson, *Functional MRI Lie Detection: Too Good to Be True?*, 36 J. AM. ACAD. PSYCHIATRY & L. 491, 493 (2008) (“[H]ow well fMRI lie detection would work in real-life situations remains an open question.”); Sean A. Spence, *Playing Devil’s Advocate: The Case Against fMRI Lie Detection*, 13 LEGAL & CRIMINOLOGICAL PSYCH. 11, 11 (2008) (suggesting that fMRI-based lie detection is inapplicable to the “real world” and lacks scientific reliability because no fMRI-based lie detection study has been replicated).

⁴⁷ For a discussion and definitions of scientific reliability and validity, albeit not identical to the usages here and in some of the relevant scientific literature, see JOHN MONAHAN & LAURENS WALKER, *SOCIAL SCIENCE IN LAW* 57–67 (6th ed. 2006).

⁴⁸ Greely & Illes, *supra* note 39, at 403–04 (noting the “artificiality of the deceptive tasks” in fMRI studies and claiming that they bear no resemblance to lying in the real world). A similar argument is that “[r]eports of finding brain patterns of activation corresponding to ‘deception’ almost always use subjects (often university students) who are told to lie about something (usually a relatively unimportant matter). Equating the lies told in such an artificial setting to the kinds of lies people tell in reality is pure fantasy at this point.” *Deceiving the Law*, *supra* note 44, at 1231.

⁴⁹ See, e.g., Kanwisher, *supra* note 41, at 12 (noting that a laboratory subject who lies because he is “instructed to do so” is not lying); Greely & Illes, *supra* note 39, at 403–04.

⁵⁰ Greely & Illes, *supra* note 39, at 402–04.

⁵¹ Critics claim that subjects can foil fMRI readings through detectable countermeasures, such as moving their tongues around, and undetectable countermeasures, such as performing simple mental arithmetic. See Kanwisher, *supra* note 41, at 12. See also Greely & Illes, *supra* note 39, at 404–05.

ing scientific outcomes.⁵² Moreover, quite a few of the experiments—indeed, most of them—have been conducted by researchers whose connection with No Lie MRI or Cephos gives them a commercial interest in the outcome.⁵³ Finally, the alleged degree of accuracy—as high as 90%, according to some claims⁵⁴—of neural lie detection is considerably higher than what could likely be expected in nonlaboratory settings.⁵⁵

Insofar as the proponents of neural lie detection have maintained their claims about the accuracy of their methods are scientifically sound and the product of scientifically valid experimentation, they appear to have been exposed as relying on flawed science. Without better evidence of external validity, without dealing with the construct validity problem of distinguishing the genuine lie from following an instruction to utter words that are not literally true, without more rigorous scrutiny of claims of reliability, without higher verified rates of accuracy, without replication, and without subjecting the research to peer review by financially disinterested scientists, the claimed ability of fMRI to identify liars appears to be just that—a claim—and far from what good scientists take to be a sound scientific conclusion.

III

ON THE RELATIONSHIP BETWEEN SCIENTIFIC AND LEGAL STANDARDS

That the science to date appears both methodologically flawed and unconvincing in its conclusions is far from the end of the story, the arguments of the skeptics notwithstanding. But the rest of the story is not a story about science. Instead, it is a story about law⁵⁶ and about the reasons for doubting that the scientific failings of fMRI-based lie detection are, or should be, dispositive for the legal system.

⁵² Kanwisher, *supra* note 41, at 13.

⁵³ For example, Christos Davatzikos, the lead researcher of the Davatzikos study, *supra* note 31, serves on the Science Board of No Lie MRI. See No Lie MRI Scientific Board, <http://noliemri.com/aboutUs/ScienceBoard.htm> (last visited Apr. 10 2010). Similarly, Frank Kozel, the lead researcher of the three Kozel studies, *supra* note 31, serves as a scientific advisor for Cephos. See Cephos Scientific Advisors, <http://cephoscorp.com/about-us/index.php#scientific> (last visited Apr. 10, 2010). Finally, No Lie MRI uses technology and methods under a license from Daniel Langleben, lead researcher on many other studies, *supra* note 31. See No Lie MRI Patents, <http://www.noliemri.com/investors/Patents.htm> (last visited Apr. 10, 2010).

⁵⁴ *Detecting Deception*, *supra* note 31, at 610. See also Davatzikos, *supra* note 31, at 663 (88%); *Telling Truth*, *supra* note 31, at 262 (78%).

⁵⁵ Henry T. Greely, *Neuroscience-Based Lie Detection: The Need for Regulation*, in USING IMAGING TO IDENTIFY DECEIT, *supra* note 41, at 46, 51 (arguing that reported accuracy rates should be “taken with a grain of salt” because the actual accuracy of fMRI lie detection on “diverse subjects in realistic settings, with or without countermeasures” is unknown).

⁵⁶ See Dale A. Nance, *Reliability and the Admissibility of Experts*, 34 SETON HALL L. REV. 191, 203 (2003) (urging the use of legal standards in evaluating scientific expertise).

Initially, we need to remind ourselves that law is about far more than putting criminals in jail, although that is the particular type of legal decision that motivates so much of the existing scientific criticism.⁵⁷ One of the scientists quoted above said that fMRI results would be especially unreliable if the subject believed that the results could “send him to prison.”⁵⁸ And another participant at the same symposium worried about a future in which the “police may request a warrant to search your brain.”⁵⁹ These may be legitimate worries, but their seriousness depends largely on a view of the Fifth Amendment privilege against self-incrimination that would characterize an involuntary lie-detection test of whatever kind as physical and nontestimonial,⁶⁰ an outcome that seems unlikely, albeit not impossible. Given that law enforcement authorities may not require a suspect to talk at all, it is difficult to imagine that a defendant’s statement could be subject to an involuntary neural evaluation of its accuracy. The circumstances in which an involuntary fMRI would be usable against a defendant would thus not only require a court to reject an explicit Supreme Court statement that the results of lie-detector tests are testimonial and hence encompassed by the Fifth Amendment⁶¹ but would also require the fMRI not to be used in conjunction with, or to test the validity of, any statement made by the defendant.⁶²

Yet even if a future that includes brain-scan warrants is a legitimate worry⁶³ that requires us to guard against police or prosecution use of fMRI lie detection, it does not follow that other potential courtroom and forensic uses of lie-detection technology are equally worrisome. Whereas a defendant’s negative results on an fMRI-based lie-detection exam can hardly suffice to prove guilt “beyond a reasonable doubt,” for example, it is a different issue when the question is whether a defendant could use an fMRI result to establish his innocence under that same and highly defendant-protective burden of

57 See, e.g., Kanwisher, *supra* note 41, at 12 (noting that the stakes of lie detection in the real world are “prison, or life, or life in prison”).

58 Deborah Halber, *Scientists: A Good Lie Detector Is Hard to Find*, MIT News, Feb. 12, 2007, <http://web.mit.edu/newsoffice/2007/lying.html> (quoting Nancy Kanishwer, Professor of Brain and Cognitive Sciences at MIT).

59 *Id.*

60 See *Schmerber v. California*, 384 U.S. 757, 761 (1966) (holding that the Fifth Amendment does not apply to physical, nontestimonial evidence taken from a suspect).

61 *Id.* at 763–64.

62 The Fifth Amendment issues are discussed in Benjamin Holley, *It’s All in Your Head: Neurotechnological Lie Detection and the Fourth and Fifth Amendments*, 28 DEV. MENTAL HEALTH L. 1, 14–22 (2009), and Sarah E. Stoller & Paul Root Wolpe, *Emerging Neurotechnologies for Lie Detection and the Fifth Amendment*, 33 AM. J.L. & MED. 359, 365–74 (2007).

63 A case in Mumbai, India has fueled this worry because the prosecution successfully used an involuntary fMRI scan to challenge the veracity of a criminal defendant. See *Deceiving the Law*, *supra* note 44, at 1231.

proof.⁶⁴ Suppose, attaching some arbitrary but conservative numbers to the existing research, that an fMRI evaluation of a defendant's claim of innocence—"I was somewhere else" or "He started the fight"—has an accuracy rate of 60 percent.⁶⁵ It is of course clear that we should not imprison people on a 60 percent probability of their guilt, and we do not do so. But the question is not whether to imprison people who are 60 percent likely to be guilty. At least that is not the only question. Equally important is whether, if there is a 60 percent chance that a defendant's claim of innocence is accurate, we would want to conclude that his guilt has been proved beyond a reasonable doubt? Indeed, this was precisely the issue in the 1998 Supreme Court case of *United States v. Scheffer*.⁶⁶ The defendant in *Scheffer* sought to introduce a polygraph test supporting the accuracy of his assertion of innocence.⁶⁷ The test results had been excluded under Rule 707 of the Military Rules of Evidence,⁶⁸ and the defendant challenged the constitutionality of Rule 707's absolute exclusion of polygraphic evidence under the Due Process⁶⁹ and Compulsory Process⁷⁰ Clauses of the Fifth and Sixth Amendments, respectively. The Supreme Court, over Justice Stevens's dissent⁷¹ and in the face of a concern about a blanket rule of exclusion on the part of four other Justices who concurred in part with Justice Thomas's opinion,⁷² held that a defendant had no constitutional right to offer polygraphic exculpatory evidence.⁷³ That the defendant may not have a constitutional right to admit a polygraph in his defense, however, does not address the nonconstitutional question of whether such evidence ought to be admissible under these or similar circumstances as a matter of policy. Moreover, the *Scheffer* majority's stark distinction be-

⁶⁴ Commentators have described the evidence that a criminal defendant needs to raise a reasonable doubt as "slight." Michael H. Graham, *Burdens of Proof and Presumptions in Criminal Cases*, 45 CRIM. L. BULL. 192 (2009).

⁶⁵ Note that the accuracy rate for identifying truth may differ from the accuracy rate for identifying a lie. Suppose a defendant claims that he was somewhere else when the crime was committed and that an fMRI indicates he is telling the truth. On the existing state of the research, this fMRI result is more reliable—has a smaller likelihood of error—than an fMRI result that indicates that the defendant's statement was false. See Kanwisher, *supra* note 41, at 11. In other words, FMRI identifies truths as lies less often than it identifies lies as truths.

⁶⁶ 523 U.S. 303 (1998).

⁶⁷ *Id.* at 306.

⁶⁸ Military Rule of Evidence 707 prohibits the admission of any lie-detection technology in military trials. *See* MIL. R. EVID. 707; *see also* *Scheffer* 523 U.S. at 306–07.

⁶⁹ U.S. CONST. amend. V.

⁷⁰ U.S. CONST. amend. VI.

⁷¹ *See* 523 U.S. at 320 (Stevens, J., dissenting).

⁷² *Id.* at 318–20 (Kennedy, J. (joined by Justices O'Connor, Ginsburg, and Breyer), concurring in part and concurring in judgment).

⁷³ *Id.* at 317.

tween “reliable” and “unreliable”⁷⁴ masks the important difference between how reliable evidence must be in order for the prosecution to use it and how reliable evidence must be in order for the defendant to use it to raise the possibility of a reasonable doubt as to his guilt, whether by buttressing his claim of innocence or, perhaps more likely, by attacking the credibility of a police officer or other prosecution witness.

Any scientific test will of course have *some* level of reliability. Whether that level of reliability is high enough for admissibility, however, depends on the purposes for which the evidence is being employed.⁷⁵ If the outcome of a test is used as the principal evidence of whether a defendant should go to prison, as it often is with DNA identification,⁷⁶ we should demand extremely high levels of reliability. But if the evidence is to be used merely as one component of a larger story about whether a defendant should go to prison, then perhaps the level of reliability can be lower— “[a] brick is not a wall,” as the famous adage in the law of evidence goes.⁷⁷ Although the standard of proof for conviction of a crime is proof beyond a reasonable doubt, it does not follow that every piece of evidence admissible to (cumulatively) establish proof beyond a reasonable doubt must be individually capable of proving beyond a reasonable doubt that the defendant is guilty.⁷⁸ To require that each piece of evidence introduced by the prosecution be reliable beyond a reasonable doubt would collapse the standard for determining guilt into the standard for determining the admissibility of an individual piece of evidence.⁷⁹ A fortiori, the level of reliability for an individual item of evidence offered as part of a larger array of evidence to show why a defendant should *not* go to

⁷⁴ *Id.* at 309.

⁷⁵ See Erica Beecher-Monas, *A Ray of Light for Judges Blinded by Science: Triers of Science and Intellectual Due Process*, 33 GA. L. REV. 1047, 1062 (1999) (arguing that reliability “is not an all-or-nothing proposition, but rather depends on the application of the evidence and the acceptable risk of error”).

⁷⁶ See, e.g., United States v. Beasley, 102 F.3d 1440, 1444–48 (8th Cir. 1996) (ruling that polymerase chain reaction DNA testing satisfied *Daubert* standards); United States v. Cuff, 37 F. Supp. 2d 279, 280–83 (S.D.N.Y. 1999) (denying motion in limine to exclude DNA evidence); State v. Bartylla, 755 N.W.2d 8, 12, 14–20 (Minn. 2008) (upholding conviction where DNA “cold hit” linked DNA profile of defendant with DNA of perpetrator that police collected at crime scene).

⁷⁷ 1 MCCORMICK ON EVIDENCE § 185, at 729 (Kenneth S. Broun ed., 6th ed. 2006). See also Judson F. Falknor, *Extrinsic Policies Affecting Admissibility*, 10 RUTGERS L. REV. 574, 576 (1956) (“[I]t is not to be supposed that every witness can make a home run.”).

⁷⁸ For more on the confusion of admissibility and sufficiency, see Dale A. Nance, *Conditional Relevance Reinterpreted*, 70 B.U. L. REV. 447, 449–59 (1990).

⁷⁹ See United States v. Glynn, 578 F. Supp. 2d 567, 574–75 (S.D.N.Y. 2008) (holding that ballistics evidence is admissible in a criminal case even if it only makes a proposition “more likely than not”); *In re Ephedra Products Liability Litigation*, 393 F. Supp. 2d 181, 187–88 (S.D.N.Y. 2005) (rejecting the argument that scientific expert testimony must be supported by “definitive scientific proof”).

prison can be lower still, arguably much lower.⁸⁰ We do not, after all, have a system in which a defendant goes to prison unless he can prove by compelling evidence or beyond a reasonable doubt that he is not guilty.

The same considerations apply to civil cases. The American legal system employs the standard of proof by a preponderance of the evidence in almost all civil cases because the failure to award damages, say, to an injured or otherwise wronged plaintiff is thought to be as serious an error as wrongly requiring a nonculpable defendant to pay damages.⁸¹ Accordingly, it is hardly clear that a party in a civil lawsuit

⁸⁰ For an extended argument in favor of asymmetry between prosecution and defense in the standards for admission of scientific evidence, see CHRISTOPHER SLOBOGIN, PROVING THE UNPROVABLE: THE ROLE OF LAW, SCIENCE, AND SPECULATION IN ADJUDICATING CULPABILITY AND DANGEROUSNESS 131–44 (2007); compare Aaron Katz, *A “Moving Bar” Approach to Assessing the Admissibility of Expert Causation Testimony*, 57 CLEV. ST. L. REV. 579, 601 (2009) (urging that the reliability requirement of *Daubert* and *Joiner* vary depending on the nature of the case in which the issue arises). The strongest response to the argument for asymmetry is that the presumption of innocence and the beyond-a-reasonable-doubt burden of proof already incorporate the justifiable, defendant-skewed epistemic goals of the criminal justice system and that overlaying special evidentiary burdens on the prosecution (or special evidentiary benefits on the defense) would be a form of double counting. See LARRY LAUDAN, TRUTH, ERROR, AND THE CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 123–28, 144 (2006). But this argument rests on the assumption, perhaps justified but perhaps not, that the existing standard of proof achieves the socially proper distribution of errors of false acquittal and false conviction. If it does not, then, given the historical provenance of the beyond-a-reasonable-doubt standard, adjusting the results of that standard through other evidentiary, substantive, or procedural devices hardly seems inappropriate. Nor is there reason to believe that the best way to achieve the optimal distribution of error is with one burden of proof rule rather than a combination of multiple evidentiary and procedural rules. See Raphael M. Goldman & Alvin I. Goldman, *Review of Truth, Error and the Criminal Law: An Essay in Legal Epistemology*, by Larry Laudan, 15 LEGAL THEORY 55, 59–60 (2009) (book review); Michael S. Pardo, *On Misshapen Stones and Criminal Law’s Epistemology*, 86 TEXAS L. REV. 347, 371–74 (2007) (reviewing LAUDAN, TRUTH, ERROR, AND THE CRIMINAL LAW, *supra*). Regardless of the outcome of the debate about asymmetry, however, the very existence of the debate and the terms on which it is conducted demonstrate the folly of trying to determine questions of the legal usability of evidence without taking legal goals and legal standards into account.

⁸¹ See *In re Winship*, 379 U.S. 358, 371 (1970) (Harlan, J., concurring) (“In a civil suit between two private parties for money damages, for example, we view it as no more serious in general for there to be an erroneous verdict in the defendant’s favor than for there to be an erroneous verdict in the plaintiff’s favor.”). On the decision-theoretic aspects of burdens of proof in civil cases in general, see James Brook, *Inevitable Errors: The Preponderance of the Evidence Standard in Civil Litigation*, 18 TULSA L.J. 79 (1982); Bruce L. Hay & Kathryn E. Spier, *Burdens of Proof in Civil Litigation: An Economic Perspective*, 26 J. LEGAL STUD. 413 (1997); John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968); Dale A. Nance, *Civility and the Burden of Proof*, 17 HARV. J.L. & PUB. POL’Y 647 (1994); Frederick Schauer & Richard Zeckhauser, *On the Degree of Confidence for Adverse Decisions*, 25 J. LEGAL STUD. 27 (1996). Ronald Allen challenges the conventional view about burdens in civil cases, arguing, correctly, that because the plaintiff must typically prove each of the multiple elements of a cause of action by a preponderance of the evidence, the actual burden on the plaintiff is substantially higher than that on the defendant. See Ronald J. Allen, *The Nature of Juridical Proof*, 13 CARDOZO L. REV. 373 (1991); Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401 (1986).

should be precluded from using lie-detection technology that is insufficiently reliable to send someone to jail to bolster his assertions about the facts of a civil case. Awarding damages is less serious than imprisoning someone, or so our legal system believes, and it is consequently a mistake to assume that a uniform standard of reliability should govern all legal uses of a particular type of evidence.

The foregoing discussion is about reliability, but the same analysis applies to questions of validity as well. The experiments that allegedly establish the reliability of fMRI lie detection have been attacked as lacking external and construct validity,⁸² but, like reliability, “scientific validity . . . is a matter of degree.”⁸³ Accordingly, whether some degree of experimental validity is good enough again depends on the use to which the experiment is being put.⁸⁴ Consider first the question of external validity—whether laboratory results permit us to draw inferences and make predictions about a different, nonlaboratory subject population.⁸⁵ This issue often arises with respect to psychological experiments when and because conclusions drawn from experiments using university undergraduates—a common pool of experimental subjects—are used to predict the behavior of non-undergraduates in nonlaboratory settings.⁸⁶ Although the population about which the predictions are made differs from the subject population, the experimental research is useful when other research demonstrates a substantial correlation between the results reached in the laboratory and those observed in nonlaboratory settings.⁸⁷ These correlations are not perfect, of course, but they are positive to a substantial degree, and whether that degree is substantial enough will depend on how the research will be used. When laboratory research is claimed to justify a policy with negative consequences for some segment of the public, for example, a higher correlation between laboratory results and nonlaboratory conclusions is necessary than when, say, the public is merely being warned to be aware of a dangerous phenomenon that thus far has been demonstrated only in the laboratory.

Although less obvious, the same considerations apply to construct validity as well. Suppose we wish to examine the relationship between eating a big breakfast and proficiency in performing mathematical tasks. Suppose also that someone were then to conduct an experi-

⁸² See *supra* notes 39–55 and accompanying text.

⁸³ Nance, *supra* note 56, at 200; See also Beecher-Monas, *supra* note 75, at 1062.

⁸⁴ See, e.g., KENNETH R. FOSTER & PETER W. HUBER, JUDGING SCIENCE: SCIENTIFIC KNOWLEDGE AND THE FEDERAL COURTS 17 (1997) (“[D]ecisions about validity depend on the needs one has for the data.”).

⁸⁵ See, e.g., Craig A. Anderson, James J. Lindsay, & Brad J. Bushman, *Research in the Psychological Laboratory: Truth or Triviality*, 8 CURRENT DIRECTIONS PSYCHOL. SCI. 3, 3 (1999).

⁸⁶ *Id.* at 7.

⁸⁷ *Id.* at 5.

ment showing a relationship between eating a big breakfast and increased ability to avoid misspellings for the rest of that day. If this experiment were used to support a claim about breakfast and mathematical proficiency, it would be open to the charge of construct validity, because what the experiment measured—spelling ability—was not the same as what it was offered to show—mathematical proficiency. But if the causes of spelling mistakes and the causes of mathematical errors were shown to be positively correlated, then an experiment showing an effect on the former would provide—in the absence of evidence of relevant differences—some evidence of an effect on the latter. The evidence would not be conclusive, but deficiencies in construct validity would not render the experiment totally spurious in terms of drawing conclusions about a different but correlated effect.

So too, perhaps, with the flaws in construct validity in many existing experiments on neural lie detection. With important exceptions,⁸⁸ the experiments that purportedly establish the reliability of fMRI lie detection are experiments in which the experimenters tell subjects to lie or not lie.⁸⁹ Critics argue that subjects are not actually lying when they follow an instruction to lie and thus that an fMRI result demonstrating a certain kind of brain activity for following an instruction to lie tells us nothing about the kinds of brain activity involved in actual lying.⁹⁰ But even though this gap between the instructed lie and the real lie poses a significant construct validity problem, it would render the experimental results valueless only if there were no correlation at all between the causes of the brain activity involved in the real lie and those involved in the instructed lie. We do not yet know whether such a correlation exists,⁹¹ but if it does, even slightly, it would again be incorrect to conclude that the existing studies offer no—as opposed to slight—support for the use of fMRI-based lie detection.

Slight support (or weak evidence) ought not to be good enough for scientists,⁹² but it is often sufficient for the law. Not only do basic

⁸⁸ See Greene & Paxton, *supra* note 31, at 12506–07 (describing an fMRI study that provided an incentive to subjects to engage in dishonesty but instructed subjects to tell the truth).

⁸⁹ See, e.g., Simpson, *supra* note 46, at 494 (“All of the published literature involves scenarios in which the volunteer subjects have been instructed to lie.”).

⁹⁰ See, e.g., Greely & Illes, *supra* note 39, at 403–04 (arguing that “[i]t is not clear how this difference [i.e., the lie instruction] from the more usual lie detection settings would affect the results” of the experiments); Kanwisher, *supra* note 41, at 12.

⁹¹ See, e.g., Simpson, *supra* note 46, at 494 (“No literature addresses the question of how this basic fact [i.e., the command to lie] affects brain activation patterns, in comparison with the more realistic situation in which the person being tested makes a completely free decision about whether to lie”).

⁹² See David H. Kaye, *Statistical Significance and the Burden of Persuasion*, 46 LAW & CONTEMP. PROBS. 13, 20 (1983) (“Typically, a scientist will not so certify evidence unless the probability of error, by standard statistical measurement, is less than 5%.”) (quoting Ethyl

principles of evidence law (as well as human thinking) routinely allow the accumulation of weak (but not spurious) pieces of evidence—whether in holistic story-creation form,⁹³ or for the related purpose of prompting an inference to the best explanation,⁹⁴ or in more linear, Bayesian fashion⁹⁵—and not only might weak evidence be sufficient to allow a defendant to resist a prosecution’s claim to have established guilt beyond a reasonable doubt, but low standards of proof pervade the legal system. In some states a plaintiff can resist a defendant’s motion for a directed verdict by offering only a “scintilla” of evidence.⁹⁶ In many contexts, evidence that is “substantial” but less than a preponderance can generate legal results.⁹⁷ And the police may stop and frisk a person upon “reasonable suspicion”⁹⁸ and obtain a search warrant by showing “probable cause” to believe that the search will yield usable evidence.⁹⁹ For these and many other purposes, weak (and thus potentially flawed) evidence serves important functions in law. Requiring highly valid scientific processes to certify evidence as “compelling,” “conclusive,” or even “highly reliable” in order for that evidence to be usable would dramatically revamp the legal system as we know it.

Corp. v. EPA, 541 F.2d 1, 28 n.58 (D.C. Cir. 1975)). See also Raichle, *supra* note 42, at 5 (equating scientific “validity” with “high statistical quality”).

⁹³ For a description of and support for the “story model,” see INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING (Reid Hastie ed., 1993); Reid Hastie, *The Role of “Stories” In Civil Jury Judgments*, 32 UNIV. MICH. J.L. REFORM 227, 229–31 (1999); Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519 (1991); Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, in INSIDE THE JUROR 192–221 (Reid Hastie ed., 1993).

⁹⁴ See Amalia Amaya, *Inference to the Best Legal Explanation*, in LEGAL EVIDENCE AND PROOF: STATISTICS, STORIES, LOGIC 135, 138 (Hendrik Kaptein, Henry Prakken & Bart Verheij eds., 2009) (describing inference to the best explanation in the law “as an ‘inference to the most coherent theory of the case’”); Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 LAW & PHIL. 223, 225 (2008) (arguing that “the process of inference to the best explanation itself best explains both the macro-structure of proof at trial and the micro-level issues regarding the relevance and value of particular items of evidence”).

⁹⁵ Bayesian approaches, inherently more incremental and thus less holistic, provide “a framework for quantifying uncertainty and methods for revising uncertainty measures in the light of acquired evidence.” Stephen E. Fienberg & Mark J. Schervish, *The Relevance of Bayesian Inference for the Presentation of Statistical Evidence and for Legal Decisionmaking*, 66 B.U. L. REV. 771, 773 (1986); Alvin I. Goldman, *Quasi-Objective Bayesianism and Legal Evidence*, 42 JURIMETRICS 237 (2002).

⁹⁶ See, e.g., Hancock v. Mid-South Mgmt. Co., 673 S.E.2d 801, 803 (S.C. 2009) (“[I]n cases applying the preponderance of the evidence burden of proof, the non-moving party is only required to submit a mere scintilla of evidence in order to withstand a motion for summary judgment.”)

⁹⁷ See, e.g., De La Fuente II v. FDIC, 332 F.3d 1208, 1220 (9th Cir. 2003).

⁹⁸ Terry v. Ohio, 392 U.S. 1, 20–22 (1968).

⁹⁹ U.S. CONST. amend. IV.

IV

JUDGES, JURIES, AND THE DANGERS OF MISUSE

Against much of the foregoing, it is often argued that juries are easily influenced by misleading evidence and are inept at critically evaluating technical evidence.¹⁰⁰ As a result, the critics argue that superficially persuasive pseudoscientific evidence will have a greater effect on deliberations than it should.¹⁰¹ Because jurors cannot evaluate scientific evidence critically and cannot appropriately weigh weak but nonspurious evidence, one weakly scientific brick will turn out, it is said, to constitute the entire wall for most jurors.

This reliance on juror incompetence to justify excluding neuroscience evidence seems misplaced, however, or, at the very least, premature. As a preliminary matter, it is worth noting that jury trials are a small and diminishing proportion of all trials,¹⁰² and we might thus want to be cautious about taking the jury as more central to legal decision making than it really is. More importantly, however, the empirical evidence on jury overvaluation is decidedly mixed.¹⁰³ Indeed, if we (and the neuroscientists) subjected the common claims of jury overvaluation to the same scrutiny that we subject scientific evidence,

¹⁰⁰ For an example of an argument that misleading evidence persuades juries, see Walter Sinnott-Armstrong et al., *Brain Images as Legal Evidence*, 5 EPISTEME 359, 367–70 (2008). On juror incompetence more generally, see Michael S. Jacobs, *Testing the Assumptions Underlying the Debate About Scientific Evidence: A Closer Look at Juror “Incompetence” and Scientific “Objectivity,”* 25 CONN. L. REV. 1083, 1086–88 (1993).

¹⁰¹ See, e.g., Gazzaniga, *supra* note 46, at 413 (cautioning against the “undisciplined” use of neuroscience evidence since jurors tend to “over-accept[]” such findings “and even prematurely grant[] the status of sheer truth to some”).

¹⁰² See Frederick Schauer, *On the Supposed Jury-Dependence of Evidence Law*, 155 U. PA. L. REV. 165, 172–75 (2006).

¹⁰³ See LAUDAN, *supra* note 80, at 214–15. For good summaries of the existing primary research, much of which suggests that juries are not nearly as inept at evaluating scientific or expert evidence as is often supposed, see NEIL VIDMAR & VALERIE P. HANS, AMERICAN JURIES: THE VERDICT 177–80 (2007); Richard D. Friedman, *Minimizing the Jury Over-Valuation Concern*, 2003 MICH. ST. L. REV. 967, 982–86 (2003); Jacobs, *supra* note 100, at 1086–93; Daniel A. Krauss & Bruce D. Sales, *The Effects of Clinical and Scientific Expert Testimony on Juror Decision Making in Capital Sentencing*, 7 PSYCHOL. PUB. POL’Y & L. 267, 273–77 (2001); Richard O. Lempert, *Civil Juries and Complex Cases: Taking Stock After Twelve Years*, in VERDICT: ASSESSING THE CIVIL JURY SYSTEM 181, 230–35 (Robert E. Litan ed. 1993); Dale A. Nance & Scott B. Morris, *An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Large and Quantifiable Random Match Probability*, 42 JURIMETRICS 403, 404 (2002); Neil Vidmar & Shari Seidman Diamond, *Juries and Expert Evidence*, 66 BROOK. L. REV. 1121, 1149–66 (2001). See also Cheryl Boudreau & Mathew D. McCubbins, *Competition in the Courtroom: When Does Expert Testimony Improve Jurors’ Decisions?*, 6 J. EMPIRICAL LEGAL STUD. 793, 814–15 (2009) (finding that juror comprehension of expert testimony can increase when opposing experts exchange reasons with each other); Joseph Sanders, *The Merits of the Paternalistic Justification for Restrictions on the Admissibility of Expert Evidence*, 33 SETON HALL L. REV. 881, 908–10, 936–38 (2003) (endorsing the jury overvaluation worry but basing it more on problems of complexity of evidence rather than on jury misunderstanding of science or expertise and suggesting that even complexity may not present insurmountable problems).

we might find that the alleged basis for excluding bad scientific evidence itself rests on less than ideal science. Consider the research purportedly showing that people take brain scan images as having more evidentiary value than such images actually have.¹⁰⁴ One study¹⁰⁵ compared the effect of textual “neurobabble” with the effect of accurate explanations; another compared brain scans to plain text, simple color bar graphs, and a topographical map;¹⁰⁶ and a third compared the effects of neuroimages with psychological testimony read aloud in insanity defense cases.¹⁰⁷ But none compared the brain scan or neural explanation to otherwise identical or even substantially similar non-brain evidence. By failing to exclude the potentially confounding variables of image type, especially photographic representation, the researchers cannot properly conclude that the distortion of evidentiary valuation was an effect of the brain scans rather than the effect of a photographic image,¹⁰⁸ of a representational image (or even drawing) in complex color, or of complex information presented without opposing explanations and opportunity for cross-examination. Consequently, we lack evidence that judges and juries overvalue brain-scan evidence compared to the kind of visual evidence routinely used in trials, and we do have some evidence that jurors may understand more than we think they do.¹⁰⁹ Moreover, in practice if not in theory, the admissibility and use of some types of evidence may vary with whether it is the judge or jury who is serving as

¹⁰⁴ See David P. McCabe & Alan D. Castel, *Seeing is Believing: The Effect of Brain Images on Judgments of Scientific Reasoning*, 107 COGNITION 343, 349–51 (2008) (concluding that “there is, indeed, something special about the brain images with respect to influencing judgments of scientific credibility”); Deena Skolnick Weisberg et al., *The Seductive Allure of Neuroscience Explanations*, 20 J. COGNITIVE NEUROSCIENCE 470, 475–77 (2008) (finding that “logically irrelevant neuroscience information can be seductive—it can have much more of an impact on participants’ judgments than it ought to”).

¹⁰⁵ See Weisberg et al., *supra* note 104, at 471–72, 475–77. The term “neurobabble” comes from Sinnott-Armstrong et al., *supra* note 100, at 368.

¹⁰⁶ McCabe & Castel, *supra* note 104, at 345–47, 349–51 (“The use of brain images to represent the level of brain activity associated with cognitive processes influenced ratings of the scientific merit of the reported research, compared to identical articles including no image, a bar graph, or a topographical map.”).

¹⁰⁷ Jessica R. Gurley & David K. Marcus, *The Effects of Neuroimaging and Brain Injury on Insanity Defenses*, 26 BEHAV. SCI. & L. 85, 93–95 (2008) (“The addition of neuroimages showing brain damage increased the likelihood of a [not guilty by reason of insanity] verdict.”).

¹⁰⁸ See Adina L. Roskies, *Are Neuroimages Like Photographs of the Brain?*, 74 PHIL. SCI. 860, 868 (2007) (criticizing the conflation of photography and neuroimaging); Sinnott-Armstrong et al., *supra* note 100, at 367–68 (discussing the impact of photographic evidence). On the distorting effect of photographs generally, see David A. Bright & Jane Goodman-Delahunty, *Gruesome Evidence and Emotion: Anger, Blame, and Jury Decision-Making*, 30 LAW & HUM. BEHAV. 183 (2006), and on the distorting effect of colored images, see Aura Hanna & Roger Remington, *The Representation of Color and Form in Long-Term Memory*, 24 MEMORY & COGNITION 322, 328–29 (1996).

¹⁰⁹ See *supra* note 103.

the trier of fact, just as it does in the case of hearsay¹¹⁰ Even to the extent that juror misperception is a legitimate worry, extrapolating this worry to the legal system generally would be mistaken, precisely because juries make none of the decisions regarding reasonable suspicion to stop, probable cause to search, and other decisions as to which the credibility of a police officer is especially at issue and because juries make only a small percentage of trial decisions.¹¹¹ Admittedly, designing an evidentiary system in which admissibility varied depending on whether the trier of fact was judge or jury would encounter difficulties, and formally, if not informally, the American legal system has rejected such an approach. But whether skepticism about juror comprehension, even if well-grounded, is the appropriate model for all of law again cannot be determined without regard to the normative goals of the legal system.

A related objection might be that allowing thresholds of reliability to vary depending on their use would be undercut by cognitive (and even precedential) contamination across those uses. Would allowing defendants to use fMRI lie-detection evidence to support a claim of innocence lead to allowing prosecutors or plaintiffs to do the same in proceeding against a possibly nonculpable defendant?¹¹² Would authorizing judges to hear fMRI lie-detection evidence in evaluating the credibility of a police officer at a suppression hearing lead to permitting juries to hear such evidence in determining ultimate guilt or innocence? We should not dismiss these potential worries as completely fanciful, but again these are empirical and causal claims about the effect of one action on another. It is more than a bit ironic that those most insistent about finding a sound scientific and empirical basis for the admission of various forms of evidence often seem comfortable abandoning science in favor of their own hunches when the question is about the potential downstream dangers of allowing certain forms of evidence to be used for a particular purpose. Those dangers may exist, but no scientific evidence exists to support such a view. For now, the empirical support for the belief that allowing fMRI lie detection by a defendant in a criminal case will lead to allowing fMRI lie detection by the prosecution against an unwilling defendant appears to be no stronger than the empirical support for the view that fMRI lie detection can actually distinguish liars from truth tellers.

¹¹⁰ See Schauer, *supra* note 102, at 166 n.3.

¹¹¹ See *id.* at 172 & n.31.

¹¹² Justice Stevens notes and responds to this objection in *United States v. Scheffer*, 523 U.S. 303, 338 n.29 (1997) (Stevens, J., dissenting) (arguing against the proposition that allowing the defendant to admit exculpatory polygraph evidence would result in permitting the prosecution “to introduce inculpatory test results”).

V
COMPARED TO WHAT?

In law, as in science, a crucial question is, “compared to what?” And this question can usefully be applied to determining witness veracity in courts of law. Traditionally, the legal system has left the assessment of witness credibility and veracity to the scientifically-unaided determination of the trier of fact, but just what mechanisms do judges and juries use to make these determinations? We know that jurors often use characteristics other than the content of what a witness says to evaluate the truth of a witness’s claim, including factors such as whether a witness looks up or down, fidgets, speaks slowly or quickly, and speaks with apparent confidence; and we know that such factors are at best highly unreliable and at worst random.¹¹³ Indeed, numerous studies of the ability of untrained people to determine truth telling in others rarely rises above 60 percent, where 50 percent is the probability that a purely random guess is correct.¹¹⁴ Moreover, the rules of evidence themselves exacerbate the problem of unreliable determination of veracity by judges and juries by presuming on the basis of scarcely more than venerable superstition that those who have been convicted of serious crimes, even crimes not involving dishonest statements, are more likely to lie than those who have not,¹¹⁵ by allowing witnesses to offer testimony about whether other witnesses have a reputation in the community for lying or truth telling,¹¹⁶ and by permitting witnesses to offer their personal opinions about the general credibility of other witnesses.¹¹⁷

We can now reframe our question. The question is not, or at least not only, whether fMRI-based lie detection is reliable enough in the abstract to be used in court. Rather, it is whether there are sound reasons to prohibit the use of evidence of witness veracity that is likely better, and is at least no worse, than the evidence of witness veracity that now dominates the litigation process. The choice is not between very good evidence of veracity and inferior fMRI evidence; it is between less good—bad, if you will—fMRI evidence and the even worse evidence that is not only permitted, but also forms the core of the common law trial process. And although it is sometimes a weak argument for a conclusion that something else is worse, if the something

¹¹³ See *supra* notes 18–19.

¹¹⁴ See Charles F. Bond, Jr. & Bella M. DePaulo, *Accuracy of Deception Judgments*, 10 J. PERSONALITY & SOC. PSYCH. REV. 214, 214, 217 (2006); see also Maureen O’Sullivan, *Why Most People Parse Palters, Fibs, Lies, Whoppers, and Other Deceptions Poorly*, in DECEPTION, *supra* note 10, at 74; Aldert Vrij et al., *Increasing Cognitive Load to Facilitate Lie Detection: The Benefit of Recalling an Event in Reverse Order*, 32 LAW & HUM. BEHAV. 253, 253 (2008).

¹¹⁵ FED. R. EVID. 609(a)(1).

¹¹⁶ FED. R. EVID. 608(a).

¹¹⁷ *Id.*

else is unlikely to change and the improvement is plausible, then the argument is not so weak after all.

VI

ON THE USES AND LIMITATIONS OF SCIENCE: DOUBTING DAUBERT

The tone of the foregoing notwithstanding, it is decidedly not my goal here to argue for the admissibility of fMRI-based lie-detection evidence in the courtroom or for related forensic purposes. Rather, it is to suggest that the reliability and validity standards for scientific evidence that courts use must be standards that come, ultimately, from the legal goals of legal institutions and not from the scientific goals of scientific institutions.¹¹⁸ Science can tell us that a certain scientific process has, say, a 12 percent error rate (or specific rates of Type I and Type II errors or false positives and false negatives). And scientists must decide for their own scientific purposes whether such rates are sufficient, for example, to assert that something is the case, conclude that a finding is adequate for publication, or find a research program promising enough to renew a research grant. But whether such an error rate is sufficient for a trier of fact to hear it, put someone in jail, keep someone out of jail, justify an injunction, or award damages is not itself a scientific question.

The same applies to methods of inquiry. Science properly relies on peer review, replication, and other indicia of sound methodology. But whether these are the right indicia for purposes of nonscientific action, including but not limited to courtroom verdicts, is not a scientific determination,¹¹⁹ and to think otherwise is often to believe erroneously that one can derive a legal or policy ought from a scientific is.

Evidence cognoscenti will detect in this a challenge to *Daubert* itself, and that may be so. *Daubert* and its successors¹²⁰ were aimed primarily at products liability and mass tort verdicts based on allegedly persuasive but unreliable junk science.¹²¹ Indeed, it is difficult to read the description of the tire failure expert in *Kumho Tire Co., Ltd. v. Carmichael*¹²² without recognizing that junk science really does ex-

¹¹⁸ See sources cited *supra* note 9. See also Margaret A. Berger, *Upsetting the Balance Between Adverse Interests: The Impact of the Supreme Court's Trilogy on Expert Testimony in Toxic Tort Litigation*, 64 LAW & CONTEMP. PROBS. 289, 300–02 (2001) (distinguishing legal and scientific standards of causation).

¹¹⁹ See DAVID L. FAIGMAN ET AL., SCIENCE IN THE LAW: STANDARDS, STATISTICS AND RESEARCH ISSUES § 1-3.5.1 (2002) (determining value of scientific expert opinion “is a matter of policy, not science”).

¹²⁰ See *supra* note 24.

¹²¹ See, e.g., United States v. Starzecpyzel, 880 F. Supp. 1027, 1039 (S.D.N.Y. 1995); JOSEPH SANDERS, BENECTIN ON TRIAL: A STUDY OF MASS TORT LITIGATION 210 (1998).

¹²² 526 U.S. 137 (1999).

ist.¹²³ And the legal system no doubt must guard against a world in which experts in astrology, phrenology, and countless other bogusologies, some of which appear superficially more plausible than astrology and phrenology but have little more grounding in fact, have a place in the courtroom. Moreover, and as the recent National Academy of Sciences study documented in detail,¹²⁴ many traditionally used methods of forensic identification—bite marks, shoe prints, handwriting analysis, ballistics, tool marks, and even fingerprints¹²⁵—have less scientific backing than their proponents have claimed and less than the legal system has historically accepted.¹²⁶

Identifying the problem is thus straightforward: prior to the *Daubert* revolution, American courts admitted into evidence experts and tests purporting to demonstrate defective manufacture, causation, or identification, but which in reality were based on empirical conclusions that had no sound scientific basis as measured by the standards of science. Without *Daubert*, so the argument goes,¹²⁷ such pseudoscientific or weak scientific evidence will continue to be admitted into evidence and continue to persuade juries notwithstanding its scientific weakness. As a consequence, innocent defendants will be convicted and nonculpable tort defendants will be held liable to an unacceptable degree.

The aforesaid steps to the *Daubert* conclusion are based on empirical claims for which there is little empirical evidence. We do not know with confidence, for example, how often the admission of scientifically substandard evidence has produced erroneous verdicts. Substandard evidence could produce erroneous verdicts in cases where the following three conditions are true: where better evidence leading to the same conclusion does not accompany the substandard evi-

123 On the junk science problem generally, see PETER W. HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM (1991).

124 NATIONAL RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 127–82 (2009).

125 See, e.g., United States v. Monteiro, 407 F. Supp. 2d 351, 355 (D. Mass. 2006) (excluding ballistics testimony); United States v. Green, 405 F.2d 104, 120–22 (D. Mass. 2005) (same); cf. Robert Epstein, *Fingerprints Meet Daubert: The Myth of Fingerprint “Science” is Revealed*, 75 S. CAL. L. REV. 605 (2002) (questioning scientific basis for fingerprint matching).

126 See Michael J. Saks, *Explaining the Tension Between the Supreme Court's Embrace of Validity as the Touchstone of Admissibility of Expert Testimony and Lower Courts' (Seeming) Rejection of Same*, 5 EPISTEME 329, 330–32 (2008) (“[F]orensic science evidence stands on a foundation more of impressions and faith than of science.”); Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069, 1094–1127 (1998).

127 Indeed, Justice Stevens acknowledged the view that *Daubert* addressed the too easy acceptance of “junk science.” See *General Electric v. Joiner*, 522 U.S. 136, 153 n.6 (1997) (Stevens, J., concurring in part and dissenting in part). See also, e.g., Maxine D. Goodman, *A Hedgehog on the Witness Stand—What's the Big Idea?: The Challenges of Using Daubert to Assess Social Science and Nonscientific Testimony*, 59 AM. U. L. REV. 635, 645 & n.43 (2010) (noting that *Daubert* addressed the problem of junk science).

dence, where the substandard evidence causes the trier of fact to find for the prosecution or plaintiff where otherwise the verdict would have been different, *and* where the defendant was not in fact guilty or culpable. These three conditions may all be present in many cases, but how many cases is uncertain. It is also uncertain what proportion of all cases, verdicts, accidents, or crimes they constitute. It is thus far from clear how much of a problem erroneous judgments are, and how effective *Daubert* has been in solving it. Nor is it clear that the best solution to the problem of junk science is one that is applied at the point of admissibility, because solving or ameliorating the junk science or erroneous verdict problem by vigorous use of summary judgment, directed verdicts, and dismissal could likely achieve the same result without mistakenly importing the all-things-considered reliability of a party's entire case into the determination of the admissibility of particular pieces of evidence.¹²⁸

But even if *Daubert* has significantly reduced the number of erroneous verdicts that poor science actually causes, the "compared to what" question still looms. Bad science is worse than good science, but may not be worse than the nonscience that lurks in the heads of judges and jurors. And flawed science is hardly worse than the superstitions and urban legends that influence so much of public policymaking and legal decision making.¹²⁹ *Daubert* is based on the sound premise that a manufacturer should not be liable for damages unless there is a genuine basis for believing that some negligent act of the manufacturer actually caused injury to the user of the product, but it is important to contemplate what occurs when bad science, measured by scientific standards, is excluded from litigation. The answer to this question is unclear, but in the absence of weak but probative science, litigants may offer, and courts may admit, even more nonexpert and nonscientific evidence. After all, the American civil litigation system does not prohibit automobile tort victims from recovering unless they can show with scientific reliability that the defendant was driving negligently, nor does it prohibit defendants from offering a wide variety of nonscientific evidence to keep themselves out of prison. Requiring that science guide all legal determinations of guilt or innocence, liability or nonliability, is utopian in both the best and worst senses of

¹²⁸ See Richard D. Friedman, *Squeezing Daubert Out of the Picture*, 33 SETON HALL L. REV. 1047, 1053–56 (2003); Nance, *supra* note 56, at 252; Michael S. Pardo, *Evidence Theory and the NAS Report on Forensic Science*, 2010 UTAH L. REV. (forthcoming 2010) (manuscript at 23–24), available at <http://ssrn.com/abstract=1511719>. For the related argument that standards for the authentication of documents are better understood and determined as matters of evidentiary sufficiency, see Lawrence A. Alexander & Elaine A. Alexander, *The Authentication of Documents Requirement: Barrier to Falsehood or to Truth?*, 10 SAN DIEGO L. REV. 266, 278–79 (1973).

¹²⁹ See *supra* notes 115–17.

that word.¹³⁰ Best in the sense that such a system might achieve more justice than the one we now have; but worst in the sense that eliminating bad or flawed science from the courtroom, the legal system, and the rules of evidence would require a dismantling of the entire edifice of common law adjudication so unlikely as to produce perverse results. Attempting to make litigation more scientific by keeping out bad science while not doing anything about the ubiquitous non-science that pervades the entire system might well result in a system that is less scientific and less reliable because it keeps out somewhat poor science and while letting the really poor science sneak in through the backdoor by not calling itself science at all.

At the heart of the controversy over law's use of poor science is the justified concern of scientists to keep *their* scientific enterprise free from nonscientific taint. When science that is not ready for prime scientific time is commandeered for commercial gain, science suffers, as with the commercialization of fMRI-based lie detection. Encouraging shoddy science for legal or policy use is bad for science, and in the long run may—and this is an empirical question—hurt us all by polluting science and devaluing its public and policy use. But the tension between the worthy goals of long-term scientific integrity and the short-term value of imperfect scientific output is hardly new and hardly unique to lie detection or the law of evidence. When medical researchers performing placebo-controlled experiments reach a point at which they suspect but do not yet know with scientific confidence that a new drug will cure a fatal disease,¹³¹ they face the same moral dilemma that Dr. Martin Arrowsmith faced in Sinclair Lewis's great novel¹³² and countless real research physicians have faced before and after: whether to sacrifice science to the alleviation of immediate suffering or to sacrifice people and their health to long-term scientific integrity. The stakes with respect to fMRI-based lie detection are typically lower, but the question is the same. If flawed or commercially motivated science is usable in law, science will suffer. But if flawed or commercially motivated science is barred from the law in the name of science, law's own goals may suffer,¹³³ and the tension and tradeoffs

¹³⁰ This realization may be why *Kumho Tire Co., Ltd. v. Carmichael* made clear that although *Daubert*'s broad concept of reliability is applicable to all expert testimony, using the norms of science to evaluate reliability is only necessary when the proposed evidence or testimony purports to be scientific. See 526 U.S. 137, 151–52 (1999); see also *supra* note 79 and accompanying text.

¹³¹ See ROBERT J. LEVINE, THE ETHICS AND REGULATION OF CLINICAL RESEARCH 185–213 (2d ed. 1986); Sharona Hoffman, *The Use of Placebos in Clinical Trials: Responsible Research or Unethical Practice?*, 33 CONN. L. REV. 449, 451 (2001).

¹³² SINCLAIR LEWIS, ARROWSMITH (1925).

¹³³ The obligation of law simply to reach a decision and the inability of law to postpone a judgment until better evidence is available are especially important in this context.

between the goals of law and of science can never be completely eliminated.¹³⁴

That the evaluation of science within the legal system must be based on characteristically legal goals, standards, and norms is another example of the partial distinctiveness of legal thinking, analysis, and decision making.¹³⁵ Judges base their decisions on *stare decisis*, but elementary textbooks on informal logic treat arguments from past practice as a fallacy. Lawyers are expected to rely on authority, but thoughtful scientists recognize that reliance on scientific authority is often at odds with scientific method. When Blackstone observed that “it is better that ten guilty persons escape, than that one innocent suffer,”¹³⁶ he not only drew on ideas now associated with Type I and Type II errors but also made clear that law’s own goals required subjugating maximum accuracy to the greater value of personal liberty.¹³⁷ As these examples show, it is a mistake to assume that the job of law is to enforce or replicate the decision-making modes of other disciplines and other domains. And so it is with the standards that law uses to determine whether evidence, scientific and otherwise, is sufficiently reliable to be usable for this or that legal purpose. Law must listen to what neuroscientists say about neuroscience, but it must also be attentive to the adjectives and adverbs. When neuroscientists say that there is no “compelling” evidence of fMRI’s lie-detecting reliability, that

See Neil B. Cohen, *The Gatekeeping Role in Civil Litigation and the Abdication of Legal Values in Favor of Scientific Values*, 33 SETON HALL L. REV. 943, 947–48 (2003).

134 When I suggest that the decision about the use or nonuse of neuroscience-based lie detection for trial or forensic purposes must be made according to legal standards, I do not also mean to suggest that the decision should be made solely by lawyers and judges. Committees or other decision-making processes which represent both legal and scientific professionals would be preferable to leaving the decision solely to legal professionals or solely to scientists. My principal concern in this paper is to argue against the view that only scientists applying scientific criteria should decide on the appropriate uses for science or its conclusions.

135 See FREDERICK SCHAUER, THINKING LIKE A LAWYER: A NEW INTRODUCTION TO LEGAL REASONING 211–12 (2009).

136 4 WILLIAM BLACKSTONE, COMMENTARIES *358. See also Alexander Volokh, Aside, *n Guilty Men*, 146 U. PA. L. REV. 173 (1997). Modern analyses of the Blackstonian maxim have understood, properly, that we should be interested in the consequences or utilities of four and not just two different outcomes—true convictions, false convictions, true acquittals, and false acquittals. See Alan Cullison, *Probability Analysis of Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach*, 1 U. TOL. L. REV. 538, 564–68 (1969); Richard D. Friedman, *Standards of Persuasion and the Distinction Between Fact and Law*, 86 NW. U. L. REV. 916, 938 (1992); Erik Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. DAVIS L. REV. 85, 93–96 (2002); Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 LAW & SOC. INQUIRY 95, 99–100 (1996); Lawrence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1375 (1971).

137 See David L. Faigman, *Expert Evidence in Flatland: The Geometry of a World Without Scientific Culture*, 34 SETON HALL L. REV. 255, 267 (2003) (achieving balance between false positives and false negatives is a legal question).

there is “very little basis” for confidence in the results produced so far, or that claims about fMRI results have been made “prematurely,” they are imposing an evaluative standard on the experimental results. This is as it should be, for we cannot make sense of these or any results without having *some* evaluative standard. But the evaluative standard to be used by the law, even when it is science that is being evaluated, must be based on law’s goals, law’s purposes, and law’s structures, and as is so often the case, what is good outside of law may not be good enough inside it. Less obvious but often more important is the corollary—that what is not good enough elsewhere may sometimes be good enough for law.

