September, 2012

Using Brain Imaging for Lie Detection: Where Science, Law, and Policy Collide

Daniel D. Langleben
Jane Campbell Moriarty, *Duquesne University School of Law*

Duquesne University
School of Law

Legal Studies Research Paper Series

Using Brain Imaging for Lie Detection: Where Science, Law, and Policy Collide

Daniel D. Langleben
Jane Campbell Moriarty

2012

Duquesne University School of Law Research Paper
No. 2012-12

Forthcoming article in Psychology, Public Policy, and Law

Psychology, Public Policy, and Law

1076-8971/12/$12.00 DOI: 10.1037/a0028841

Using Brain Imaging for Lie Detection: Where Science, Law, and Policy Collide

Daniel D. Langleben
University of Pennsylvania

Jane Campbell Moriarty
Duquesne University

Progress in the use of functional magnetic resonance imaging (fMRI) of the brain to differentiate lying from truth-telling has created an expectation of a breakthrough in the search for objective methods of lie detection. In the last few years, litigants have attempted to introduce fMRI-based lie detection evidence in courts. Both the science and its possible use as courtroom evidence have spawned much scholarly discussion. This article contributes to the interdisciplinary debate by identifying the missing pieces of the scientific puzzle that need to be completed if fMRI-based lie detection is to meet the standards of either legal reliability or general acceptance. The article provides a balanced analysis of the current science and the cases in which litigants have sought to introduce fMRI-based lie detection. Identifying the key limitations of the science as expert evidence, the article explores the problems that arise from using scientific evidence before it is proven valid and reliable. We conclude that the *Daubert*'s "known error rate" is the key concept linking the legal and scientific standards. We suggest that properly controlled clinical trials are the most convincing means to confirm or disprove the relevance of this promising laboratory research. Given the controversial nature and potential societal impact of this technology, collaboration of several government agencies may be required to sponsor impartial and comprehensive clinical trials that will guide the development of forensic fMRI technology.

AQ: 1

Keywords: functional magnetic resonance imaging (fMRI), brain, deception, lie detection, neuroscience

Recent progress in the use of functional magnetic resonance imaging (fMRI) of the brain to evaluate deception and differentiate lying and truth-telling has created anticipation of a breakthrough in the search for technology-based methods of lie detection. Attempts by commercial entities to introduce fMRI lie detection evidence in courts have prompted commentary and criticism on both ethical and scientific grounds without a corresponding generation of new research data to address such concerns.

Major unanswered questions include the sensitivity of the new technology to countermeasures, its external validity and accuracy, and the specificity of the observed fMRI patterns to deception. Our review suggests that although these are important, the critical knot of law and science that must be untangled to permit further translational progress is the determination of the "error rates" of the technology as defined by the *Daubert* criteria of admissibility. This determination includes not only the accuracy of tests within each subject, but also their predictive power in the relevant population.

Daniel D. Langleben, Perelman School of Medicine, University of Pennsylvania; and Jane Campbell Moriarty, School of Law, Duquesne University.

We thank Kenneth Foster of the University of Pennsylvania for helpful comments on an early draft; Christy Gamble and Gabrielle Steier of the Duquesne University School of Law for research assistance; and Bridget Praskovich for editorial and formatting assistance.

Correspondence concerning this article should be addressed to Jane Campbell Moriarty, Carol Los Mansmann Chair in Faculty Scholarship, Professor of Law and Associate Dean for Faculty Scholarship, Duquesne University School of Law, 600 Forbes Avenue, Pittsburgh, PA 15282.
AQ: 57 E-mail: moriartyj@duq.edu

The article seeks to explain for the interdisciplinary audience the pivotal difference between small-scale experimental research studies and properly controlled clinical trials. We emphasize that such trials are critical to evidentiary reliability. Prior to such trials, expert testimony that a given witness is deceptive in response to a given question remains a risky leap from existing data. Given the AQ: 2 multidisciplinary nature of the research and the diversity of special interests involved, funding clinical trials of fMRI-based lie detection technology is not a trivial endeavor. In light of its potential importance to society and the fields of law and medicine, we propose a public funding initiative leading to a peer-reviewed translational research program on the brain mechanisms of deception with a special emphasis on multicenter clinical trials of fMRI-based lie detection.

The perils of admitting unproven scientific evidence in trials are well known, as discussed in the National Research Council of the National Academies' report (NRC Report), *Strengthening Forensic Science in the United States: A Path Forward* (2009). The Report AQ: 3 criticizes many forms of forensic science, such as handwriting and fingerprint comparisons, arson, and bitemark evidence, and notes the relationship between unproven forensic science and wrongful conviction. DNA analysis has uncovered "a disturbing number of wrongful convictions–some for capital crimes–and exposing serious limitations in some of the forensic science approaches commonly used in the United States." (NRC Report, 2009, p. 42). Nonetheless, courts continue to admit various types of forensic science evidence, frequently ignoring its documented shortcomings. Once admitted, scientific evidence tends to become rooted and difficult to eradicate later. We believe this report should encourage the legal community to require that the emerging field of forensic neuroimaging, including fMRI-based lie detection, have a proper scientific foundation before being admitted in courts.

The Development of Lie Detection Technology

The U.S. judicial system places great weight in the belief that juries are effective and reliable in determining the credibility of the witness. Yet, behavioral and social research has shown that humans are good at lying and quite poor at lie detection (Vrij, 2008). For example, an average person's ability to detect deception in a face-to-face interaction with another individual is only modestly better than chance (Ekman & O'Sullivan, 1991). Thus, the critical importance of truthful testimony and the inadequacy of human lie detectors have prompted the perennial search for a technology-based, objective method of lie detection or truth verification; this search continues today (Grubin, 2010; NRC Report, 2009; Stern, 2003).

The polygraph, which measures activity of the peripheral nervous system to detect deception, has been the primary technical method for lie detection during the last century. Beginning with the *Frye v. United States* (1923) decision, most U.S. courts have expressed disapproval of polygraph-based evidence. The U.S. Supreme Court has noted the lack of consensus on the reliability and admissibility of the polygraph (*United States v. Scheffer*, 1998) and courts remain largely hostile to its admission into evidence (Faigman, Kaye, Saks, & Sanders, 2010; Gallini, 2010). A meta-analysis commissioned by the U.S. Department of Defense found the sensitivity and specificity of the polygraph to be 59 and 92%, respectively (Crewson, 2001). The National Academy of Sciences report (Stern, 2003) laments the lack of definitive research on the accuracy of the polygraph under various conditions and estimates it to be in the vicinity of 75%; as high as 99% and as low as 55% depending on the setting (i.e., experimental vs. forensic), questioning format, the operator, and response classification rules.

The polygraph is still widely used outside the courtroom in the United States, in particular, as pre-employment and in-employment screening for some government agencies, such as the U.S. Federal Bureau of Investigation. Anecdotal evidence (U.S. Senate, 1994) and some retrospective studies have led many scholars to believe that the polygraph would perform poorly in this capacity. Due to the relatively low prevalence of the types of misconduct targeted by polygraph examinations among the U.S. government workers, most individuals flagged by the polygraph are likely to be false-positives, and a substantial proportion of the liars are likely to be missed (Baldessarini, Finklestein, & Arana, 1983; Raichle, 2009; Wolpe, Foster, & Langleben, 2005).

The more recently developed physiological measures considered to have potential for lie detection are electroencephalography (EEG) and fMRI. Both are established medical technologies developed and widely used for the assessment of brain activity. The EEG dates back to the 1920s (Berger, 1929), while a BOLD fMRI study of regional brain activity in humans was first reported in 1992 (Kwong 1992). The two techniques critically differ from the polygraph in that they measure the central (brain) rather than the peripheral (galvanic skin response, heart rate, blood pressure, and respiration) correlates of the nervous system activity. EEG-based lie detection was pioneered by Rosenfeld, (Rosenfeld et al., 1988), and has been a topic of sustained research since. fMRI is greatly superior to EEG in its ability to localize the source of the signal in the brain. EEG, on the other hand, is significantly less expensive, more mobile, and has a better time resolution than fMRI. Recent progress in the ability of fMRI to reliably measure and localize the activity of the central nervous system has created the expectation that an fMRI-based system would be superior to both the polygraph and the EEG for lie detection.

The Basic Science

The scientific and forensic concerns of fMRI-based lie detection are reviewed in greater detail elsewhere (Langleben, 2008; Langleben, Willard, & Moriarty, 2012; Spence, 2008), so we provide only a basic overview here.

BOLD fMRI and the Principle of Cognitive Subtraction

MRI is a medical imaging technique using high magnetic fields and nonionizing electromagnetic radiation to produce high-resolution, three-dimensional (3D) tomographic images of the body (Lauterbur, 1973). fMRI is distinguished from regular (structural) MRI by the speed of acquisition of each 3D image. In fMRI, serial images of the entire brain are acquired every few seconds, which is fast enough to observe changes in the regional blood volume and flow that are associated with cognitive activity. Blood oxygenation level−dependent (BOLD) imaging is presently the fMRI technique most commonly used in cognitive neuroscience (Kwong et al., 1992). BOLD relies on the difference in the magnetic properties of the contents of the blood vessels and the surrounding brain tissue as well as the magnetic difference between oxygenated and deoxygenated hemoglobin (Gjedde, 2001). BOLD fMRI does not depict absolute regional brain activity; rather, it indicates relative changes in regional activity over time. To make inferences about the nature of the regional brain activity, BOLD fMRI task designs rely on a principle of "cognitive subtraction" (Aguirre & D'Esposito, 1999). This principle assumes that the fMRI signal difference between two behavioral conditions that are identical in all but a single variable is due to this variable. Therefore, a proper comparison (i.e., control) condition is critical for meaningful BOLD fMRI paradigm (Gjedde, 2001). The fMRI activation maps reported in the literature usually represent a statistical subtraction between the fMRI activity maps related to the target and control variables (Owen, Epstein, & Johnsrude, 2001). Ideally, the comparison and target conditions would be identical, except for a single variable of interest. For example, statistically comparing the fMRI signal acquired when looking at a random sequence of white and black squares of the same size would yield the difference between brain processing of the colors white and black (Owen et al., 2001). In an fMRI deception experiment, questions that could invoke a lie or a truth could be substituted for the two types of squares, but the same principle applies.

Experiment Design

fMRI deception experiments have several critical parameters, some of which are unique to fMRI, and others that have been developed in basic psychological and polygraph research (Miller & Stiff, 1993). The *scenario* of a deception task refers to the hypothetical setting in which experimental deception takes place.

For example, some experiments involve participants in a mock crime situation and then question them about it (Kozel et al., 2005). Others probe participants about autobiographical information of different levels of intimacy (Abe et al., 2009; Spence et al., 2001). Finally, experiments that treated emotion, embarrassment, and autobiographical memory as confounds, rather than variables of interest, used relatively "neutral" scenarios that required concealing possession of a playing card for a monetary reward (Langleben et al., 2002). The task scenario also determines the risk−benefit ratio of the deception experiment. For example, critics of the practical relevance of fMRI deception research argue that the substantially lower risk−benefit ratio of deception using the concealed playing card scenario, compared to lying about an actual crime, should lead to significantly different fMRI patterns associated with deception under these two scenarios. This debate can only be resolved by direct experimental manipulation of the risk−benefit ratios of the deception experiments.

The fMRI *paradigm* refers to the way in which stimuli are presented during an fMRI scan (Donaldson & Buckner, 2001). In "event-related" paradigms, stimuli (events) and the associated brain responses are time locked to 3D fMRI images, typically acquired every 1 to 4 seconds. Event-related designs are able to characterize brain response to specific stimuli and classes of stimuli, such as possible lies and known truths (Donaldson & Buckner, 2001). However, because of their low statistical power, event-related designs require random repetition of each class (i.e., lie or truth) of stimuli up to a dozen times during an experiment. Event-related fMRI paradigms are more relevant for deception than other designs, and most of the recent deception experiments have used this approach.

The experimental deception *model* refers to the method used to generate deceptive responses and the appropriate controls. The two basic deception-generating models are the Comparison Question Test (CQT) and the Guilty Knowledge Task (GKT), also referred to as the Concealed Information Test (CIT). These models are not unique to fMRI research, and they were developed for forensic investigative use (Ben-Shakhar, Bar-Hillel, & Kremnitzer, 2002; Lykken, 1991; Stern, 2003) with the polygraph and later with EEG (Rosenfeld et al., 1988). In the CQT, test-takers answer a series of questions. One subset consists of questions unrelated to the topic of questioning, with the correct response known or presumed to be known. These questions are selected to be similar to the relevant questions in their attentional quality (e.g., salience; Raskin & Honts, 2001). The inherent subjectivity of what constitutes comparable salience makes it difficult to adequately control these questions, a main criticism for the CQT's detractors (Ben-Shakhar, 1991).

The GKT or CIT involves a series of questions designed to elicit a fixed uniform response (typically "no") to multiple items, including an item of knowledge that a "guilty" subject would seek to conceal. A negative response to such an item would constitute a forced deception that is hypothesized to have higher salience than other items (Lykken, 1991). Although not having the control problems of the CQT, the CIT's reliance on the salience of deception (rather than on the deceptive response itself) limits its specificity. The CIT is unpopular among polygraph examiners in the United States who believe that obtaining pieces of information known only to a perpetrator is often impractical. However, it is the primary model used by law enforcement in Japan, where polygraph evidence is admissible in court (Ben-Shakhar, 2001; Nakayama, 2001).

Another critical parameter of experimental deception-generating models is whether responding deceptively is being *endorsed* by the experimenter (Miller & Stiff, 1993). While in the real world an individual's deception would generally be undesirable to its target, in most deception experiments, subjects are explicitly instructed to lie (i.e., endorse) some of the questions (Spence et al., 2001). Such endorsement severely limits the ecological validity of the experiment. Some deception experiments have attempted to enhance ecological validity by allowing the subjects to choose when to lie during the task (Lee et al. 2002). Others have removed the appearance of endorsement of deception by separating the research team member who instructs participants to lie from the rest of the team, thus creating a "coconspirator" (Langleben et al., 2005).

Early Conclusions and Within-Subject Accuracy

Since 2000, academic researchers in several countries have used BOLD fMRI to study brain activity during experimental deception and malingering (Langleben et al., 2002; Lee et al., 2002; Spence et al., 2001). These early studies had to pool data from multiple subjects to make their findings. Subsequent improvement in fMRI technology permitted discrimination between an investigator-endorsed lie and truth in healthy individual subjects with an accuracy of over 75% (Davatzikos et al., 2005; Kozel et al., 2005; Langleben et al., 2005). Although there remained inconsistencies across the studies, "there has nevertheless emerged a recurrent pattern of findings suggesting that at some point in the future functional neuroimaging may be used to detect deception in situations that have significant legal consequence " (Spence et al., 2010, p. 1352). AQ: 4

Along with the experimental progress, researchers recognized and explored the limitations and the pitfalls on the way to translation of this technology to clinical use (Kozel et al., 2005). Wolpe et al. (2005) and Happel (2005) were the first to elaborate that a comprehensive understanding of the new technology's error rates requires not only the recently reported within-subject accuracy, but also the predictive power of the test, (Hyman, 2010). The predictive power combines the inherent accuracy of a test and the expected prevalence of liars in the tested population, and is a recognized milestone in the evaluation of clinical tests (Baldessarini et al., 1983).

Reactions to the Early Scientific Discoveries

After the initial fMRI studies were completed and published, the Trustees of the University of Pennsylvania and of the University of South Carolina filed separate patent applications for the technology and licensed it to start-up firms, Cephos and No Lie MRI. Articles in *The New York Times* and other publications quickly piqued the public's interest in the forensic use of fMRI technology to detect deception (Henig, 2006; Talbot, 2007).

Legal and ethics scholars also began to weigh in on fMRI lie detection (Greely & Illes, 2007; Moriarty, 2008). Criticism included the obvious technical knowledge gaps that needed to be addressed and the potential societal risks and benefits of improving lie detectors and deception research (Wolpe et al., 2005), constitutional implications (Fox, 2009; Halliburton, 2009; Pardo, 2006), AQ:5,6 and privacy concerns (Greely, 2006; Happel, 2005; Thompson, 2005). Others suggested that although validation studies were necessary for translation of fMRI lie detection into forensic practice, such studies were ethically and methodologically challenging (Halber, 2007; Kanwisher, 2009). Halber (2007) argued that the accuracy rates of 80−90%, as reported in laboratory experiments, proved the method was inadequate for field applications.

Some suggested outright regulation (Canli et al., 2007; Greely & Illes, 2007). Tovino (2007) suggested banning fMRI veracity testing outside of research until it was determined to be highly effective, and another (Moriarty, 2009) urged courts to self-impose a moratorium period to sort through the myriad scientific and jurisprudential issues at stake.

France has taken the controversial step of banning commercial use of brain imaging, but permitting its use in court. A new law, passed in 2011, provides that "brain-imaging methods can be used only for medical or scientific research purposes or in the context of court expertise" (Oullier, 2012, p. 7). According to a recent article published in *Nature*, none of the neuroscientists consulted during the drafting process encouraged the courtroom use of neuroimages (Oullier, 2012). AQ: 7

fMRI Lie Detection in Court

Despite sustained criticism by both scientific and legal scholars, for-profit companies continued to push aggressively toward the courtroom. In spring 2010, a New York State trial judge excluded fMRI expert testimony about a witness's truthfulness in *Wilson v. Corestaff Services, L. P.* (2010). A few weeks later, a federal court in Tennessee granted the government's motion to exclude fMRI expert testimony about defendant's veracity in *United States v. Semrau* (2010). In both cases, parties sought to introduce the testimony of Dr. Steven Laken, CEO of Cephos, a private company offering "lie detection/truth verification" with fMRI.

Wilson v. Corestaff Services, L. P. (2010), was an employment discrimination suit in which the plaintiff offered fMRI testimony to shore up the credibility of a main witness. The defense filed a motion *in limine* to exclude such testimony, which the trial court granted without an evidentiary hearing. The court disallowed Dr. Laken's testimony because the proposed testimony concerned a collateral matter—credibility of a witness—remarking that "anything that impinges on the province of the jury on issues of credibility should be treated with a great deal of skepticism" (p. 642). The court also held that the testimony did not meet the *Frye* standard of admissibility, which requires novel scientific evidence to be generally accepted in the field to which it belongs: AQ: 8

> Even a cursory review of the scientific literature demonstrates that the plaintiff is unable to establish that the use of the fMRI test to determine truthfulness or deceit is accepted as reliable in the relevant scientific community. The scientific literature raises serious issues about the lack of acceptance of the fMRI test in the scientific community to show a person's past mental state or to gauge credibility (p. 642). AQ: 9

There was no evidentiary hearing in *Wilson* and it settled without an appeal, so it is of marginal utility in terms of precedent. Nonetheless, for other states following the *Frye* standard, *Wilson* might be cited for its holding that the science lacks "general acceptance" in the field.

In *United States v. Semrau* (2010), the trial court held an extensive evidentiary hearing to determine whether the proposed fMRI lie detection evidence was sufficiently reliable to be admitted at trial. Dr. Laken testified that the defendant was truthful when he denied committing Medicare fraud. Dr. Laken repeated the testing session on three consecutive occasions, due to alleged problems in the first two. The first session was negative for deception. The second session was positive for deception, but Semrau complained of fatigue during the scan. Apparently, the second session had long conditional questions such as "Prior to the government's lawsuit and this indictment, were you told that the billing for AIMS testing by psychiatrists under CPT Code 99301 was inappropriate?" (*United States v. Semrau,* 2010, p. *5). The third session used reformulated test questions and was again negative for deception. Remarkably, Laken testified that he could not state whether Semrau was truthful with respect to any "specific incident question"; he could only testify to an overall picture of truthfulness. Dr. Laken may have been referring to the averaging of fMRI signals associated with multiple events of the same class (e.g. "probable lie").

Magistrate Judge Tu Pham, appointed by the federal district court judge to hear the evidentiary motion, admitted expert testimony and reviewed experts' affidavits. He analyzed the matter under substantially overlapping legal reliability standards— Federal Rule of Evidence (FRE) 702 governing expert testimony and the Supreme Court *Daubert* factors of (a) testability, (b) publication and peer review, (c) known error rate, (d) maintenance of standards and controls, and (c) general acceptance (*Daubert v. Merrell Dow Pharms., Inc.*, 1993).

The court found that the subject matter was tested and published in peer review journals, citing both legal and science journals discussing fMRI lie detection studies. Judge Pham was more troubled by Cephos' claims about its tests' error rates and testing standards. The court focused on the lack of ecological validity, remarking "[t]here are no known error rates for fMRI-based lie detection outside the laboratory setting, that is, in the 'real-world' or 'real-life' setting" (p. 11)—a concern it voiced about both AQ: 10 polygraph and fMRI lie detection.

The judge also reviewed other limitations and shortcomings of the fMRI studies that diminished the claim of a meaningful error rate: Though peer-reviewed, all studies had small ($N < 60$) samples, included young and healthy participants who were not representative of the general population, and used different types of deception-generating paradigms. Further, the court opined that the critical flaw was the difference between the motivation of the research participants and real-world suspects to lie. Finally, the court noted that all reviewed studies involved the investigators directing the participants to lie to various extents, possibly detecting brain activity related to task compliance rather than deception. In sum, the court held that, based on the current state of the science, the "real-life" error rate of fMRI-based lie detection was still unknown: a point with which we concur.

With respect to standards and controls, the court was troubled by the repeated tests used in the case at issue. The "decision to conduct a third test begs the question whether a fourth scan would have revealed Dr. Semrau to be deceptive again" (*United States v. Semrau*, 2010, p. 13). The court determined that the use of fMRI AQ: 11 for deception in the real world was not generally accepted by the scientific community, and it concluded there was insufficient proof of legal reliability of the proposed evidence.

The court also held, pursuant to FRE 403, that any probative value was substantially outweighed by the danger of unfair prejudice. By analogy to polygraph cases, the court noted that lie detection evidence to bolster credibility was highly prejudicial, particularly when credibility was a key issue and the scans were conducted without the prosecution's knowledge. In addition, the court was troubled by Dr. Laken's inability to state that Semrau was truthful as to any specific question, but could offer only a general impression of the subject's truthfulness.

Semrau was convicted and has appealed, alleging the trial court erred in excluding the fMRI evidence (*United States v. Semrau*, 2011). This case provides the U.S. Court of Appeals for the Sixth Circuit an opportunity to write an opinion with potentially precedential value for that circuit and persuasive value to other federal courts.

The Legal Implications of *Semrau* and *Wilson*

Where do these cases leave the admissibility of fMRI evidence of deception? One must be careful about inferring too much from two trial court cases, particularly as one (*Wilson*) settled without appeal and the other (*Semrau*) is still evolving. Nonetheless, we can draw some limited general conclusions that might have predictive value about the legal future of fMRI veracity evidence, and we believe that *Semrau* (and, to a lesser extent, *Wilson*) will be influential. We also address some of the competing arguments that might favor the admission of such evidence at this time.

These cases address several primary concerns that will likely be the focus of other courts' decisions as well: credibility, reliability, relevance, general acceptance, and unfair prejudice. First, both opinions focused on the subject matter of the evidence—credibility. *Wilson* held that jurors did not need expert testimony on credibility; *Semrau* (2010) echoed U.S. Supreme Court concerns that collateral litigation over lie detection "threatens to distract the jury from its central function of determining guilt or AQ: 12 innocence" (p. 15, n.21). We believe the courts will continue to be troubled by testimony that comments directly on credibility.

The jury's role as arbiter of credibility has long-standing, carefully cultivated jurisprudential roots (Fisher, 1997; Seaman, 2008), and a majority of courts have disallowed testimony that comments directly on the veracity of a particular witness, finding it not helpful to the jury or having little probative value (Kaye, Bernstein, & Mnookin, 2012; Faigman, Kaye, Saks, & Sanders, 2010−2011). "[E]xpert testimony which does nothing but vouch for the testimony of another witness encroaches upon the jury's vital and exclusive function to make credibility determinations, and therefore does not 'assist the trier of fact' as required by FRE AQ:13,14 702" (*United States v. Charley*, 1999, p. 1267). Not all courts, however, disfavor such testimony, and a minority of jurisdictions have held that the trial court has discretion to decide if expert testimony on veracity should be admitted (Kaye et al., 2012, citing cases).

There are exceptions to the prohibition against experts providing testimony that comments on credibility. For example, experts routinely testify about witnesses suffering from serious mental illnesses that may cause delusions (Melton, Petrila, Poythress, & Slobogin, 2007). Additionally, many courts have admitted expert evidence that *indirectly* comments on credibility, particularly behavioral science testimony about child sexual abuse, behaviors of battered spouses, suggestibility of children in interrogations, problems of eyewitness identification, and reasons for false confession (Faigman et al., 2010−2011; Monahan, Walker, & Mitchell, 2008; Myers, 2010; Poulin, 2007). This testimony, often termed "social framework evidence," permits experts to testify about general social science research results that are used to "construct a frame of reference or background context for deciding factual issues crucial to the resolution of a specific case" (Monahan et al., 2008, p. 1716; Walker & Monahan, 1987). Much of this testimony helps AQ: 15 the jury decide whether a given witness is credible, without specifically commenting on the truthfulness of any particular witness. Not all courts approve of social framework testimony (particularly about eyewitness identification and false confession), holding that it is not helpful to the jury in making decisions about witness credibility (*United States v. Lumpkin*, 1999). Other courts find social framework evidence too general to be helpful, because it is not about a particular witness, as noted by Monahan et al. (2008, citing cases).

Except when parties stipulate to its admissibility, most courts hold that polygraph evidence is generally inadmissible. "Throughout the 20th century, courts have been, at best, skeptical of polygraph tests and, at worst and more usual, hostile to them" (Faigman et al., 2010–2011, at § 40.1). Although such hostility may be due mostly the polygraph's limited reliability (Gallini, 2010; *United States v. Scheffer*, 1998; Stern, 2003), courts are concerned about invading the jury's role (*United States v. Swayze*, 2004) and simply may be uncomfortable with technology that purports to know when people are lying.

The U.S. Court of Appeals for the Sixth Circuit (the court deciding Semrau's appeal) has held that polygraph evidence is presumptively inadmissible in the absence of a stipulation and highly prejudicial where credibility is central to the verdict (*United States v. Sherlin*, 1995). Nonetheless, discretion is granted to the AQ: 16 trial court to decide whether the probative value of the polygraph evidence outweighs its prejudice (*United States v. Sherlin*, 1995, using a modified FRE 403 test). We believe it is likely that the AQ: 17 Sixth Circuit and many courts will react with disfavor to fMRI lie detection, reasoning that the evidence is about a collateral matter, is a direct comment on the credibility of a particular witness, and is unhelpful to the jury.

Second, many courts will focus on the reliability of the evidence. *Daubert*'s criteria, especially the "error rates" standard, are formidable, and many courts will likely find that fMRI lie detection cannot meet them at this juncture. The *Semrau* analysis is deep, careful, and compelling and will likely find traction with other courts. The experts cited and quoted in the opinion are considered well-qualified and authoritative. The current limitations of the science, as discussed in the opinion, are important concerns. As such, the next proponent of the fMRI credibility assessment evidence will find it difficult to encourage a court to disregard the findings of the *Semrau* court.

More specifically, the concerns raised in *Semrau* about the lack of ecological validity will likely be troubling for other courts assessing the evidentiary reliability. The experimental data on fMRI lie detection has been derived from small-scale laboratory studies of "normal" participants; they have not been tested either in real-life situations or in populations that deviate from what is

tapraid5/zdn-law/zdn-law/zdn00113/zdn2121d12z | xppws | S=1 | 8/23/12 | 0:18 | Art: 2011-0374

considered "normal" in experimental research. Additionally, these data were not derived from paradigms involving a level of risk to the participant that would approximate the risk−benefit ratio of deception in Semrau's case. The *Semrau* court's analysis of the shortcomings and limitation of the technology's problematic "real-world" error rate is compelling, and we anticipate that most courts using a *Daubert*-type reliability standard will be inclined to follow *Semrau*'s reasoning. Reliability must be judged on a case-by-case basis and not globally (Risinger, 2000), so it is conceivable that another litigant could make a more compelling showing in the courtroom about the reliability of fMRI lie detection. Nonetheless, it is currently difficult to separate the state of the science from any individual case.

In addition to concerns that the evidence was not sufficiently reliable, the proposed evidence in *Semrau* was not a good "fit" with the questions at issue, because the research studies could not be meaningfully applied to the truthfulness of the witness on the stand. The concept of "fit" considers whether the proposed evidence is relevant to resolving a fact in issue. In cases involving scientific evidence, *Daubert* recognizes that "scientific validity for one purpose is not necessarily scientific validity for other, unre-
AQ: 18 lated purposes" (*Daubert*, 1993, p. 591). The relevance of fMRI lie detection is inextricably tied to its reliability and FRE 702 requires a "valid scientific connection to the pertinent inquiry as a precon-
AQ: 19 dition to admissibility" (*Daubert*, 1993, p. 580). Thus, under both a relevance and a reliability analysis, fMRI evidence currently falls short of what is required for admissibility.

Fourth, courts that use the *Frye* general acceptance test (as *Wilson* did) will also likely disfavor the evidence. Although a few scientists on the advisory board of Cephos filed affidavits in support of the science, most other neuroscientists involved with the fMRI lie detection research have agreed that it is not yet ready for forensic application (Spence, 2008). Thus, without new compelling data, a party seeking to prove general acceptance will have difficulty finding credible support within the scientific community. The multifactor *Daubert* evidentiary reliability standard likewise includes general acceptance in its analysis. The lack of general acceptance among scientists in the field may well be critical to courts that follow *Daubert*.

Finally, we cannot fully discount the potential problems the combined effect of the superficial vividness of the evidence poses for fact finders unable to grasp the true scientific and statistical complexities of the fMRI technology. Early studies suggested that realistic brain images could influence the jury beyond what the
AQ: 20 evidence warrants (McCabe & Castel, 2008; Weisberg, Kiel, Goodstein, Rawson, & Gray, 2008), although there has been criticism of those studies (Schauer, 2010a). More recent data have suggested that such images are not as overwhelmingly influential to a jury as originally believed. A recent large-scale study with a meta-analysis examined the influence of neuroscience expert testimony and neuroimaging testimony on mock juries determining guilt in a criminal case in which the defendant claimed not to have requisite intent to harm the victim (Schweitzer et al., 2011). Schweitzer et al. concluded that "the overwhelming consistent finding has been a lack of any impact of neuroimages on the
AQ: 21 decisions of our mock jurors" (p. 382). In the meta-analysis, Schweitzer et al. did find that a neurological explanation for defendant's mental state—with or without brain images—was more influential to the jurors than a clinical psychological explanation. Although this study is compelling, there is more to be done in the area, a point well explained by Schweitzer et al. AQ: 22

It is likely the Sixth Circuit will affirm the lower court's decision in *Semrau,* because federal appellate courts review lower court decisions about expert evidence under an abuse of discretion standard (*General Electric Co. v. Joiner*, 1997). It is also unlikely the appellate court will find that the trial court in *Semrau* abused its discretion by excluding the proposed testimony under FRE 403. "In general, an abuse of discretion will be found only if the trial court's decision is 'arbitrary,' 'irrational,' 'capricious,' 'whimsical,' 'fanciful,' or 'unreasonable' . . . [and] . . . the . . . exercise of its discretion will not be disturbed unless it can be said that 'no reasonable person would adopt the district court's view'" (Nicolas, 2004, p. 533). The *Semrau* decision is well reasoned and well AQ:23,24,25
grounded in both facts and science, and it is unlikely an appellate court will overturn it. Even if *Semrau* is affirmed, however, the appellate court may choose not to address the issue in depth, simply finding that the lower court did not abuse its discretion. If that happens, then the magistrate judge's opinion may not carry much weight with other courts, because it may be considered an opinion limited to the facts of that case. Additionally, even if the Sixth Circuit writes an in-depth opinion on the reliability and admissibility of fMRI lie detection evidence, it will not be binding on other courts outside of the circuit, and other federal courts may disregard it. Finally, the appellate court may find that the defendant simply failed to meet the reliability standard in this case, but make no comment on the reliability of the science in general. Thus, the inadmissibility of this evidence is by no means certain in other courts. Yet, we believe the reasoning in *Semrau* will be persuasive, given the quality of the court's analysis and its detailed explanation of the current limitations of fMRI lie detection.

However, there are competing arguments that might favor admission of the testimony in future cases. Juries' subjective assessments of credibility are terribly unreliable (Schauer, 2010b). The basic fMRI veracity research is sound science of the type envisioned by *Daubert*: It is peer-reviewed research done by various scientists in quality laboratories under well-controlled conditions. If courts admit the evidence, it should be presented as probabilistic rather than a categorical conclusion that a given witness is truthful or deceptive. Such probabilistic testimony should be based on the known error rates (i.e. overall accuracy and false positive and negative rates) of the experimental studies available at the time of the testimony. By introducing the evidence with known error rates (rather than as a categorical conclusion), juries may be able to evaluate the evidence in a more balanced fashion. Empirical scholarship suggests that juries do not necessarily overvalue random match probabilities and can make reasonable use of complex material with appropriate instruction (Nance & Morris, 2002, 2005). Thus, fMRI lie detection evidence, which would present less robust statistical significance than DNA evidence, may also not be overvalued by the jury. Additionally, the fMRI veracity research is also far more experimentally grounded than the commonly admitted individualization evidence (e.g., fingerprints, handwriting, tool marks) roundly criticized by the NRC Report (2009). Finally, other forms of neuroimaging, such as nuclear medicine (positron emission tomography [PET] and single-photon emission computed tomography [SPECT]) evidence, are often admitted in civil and criminal trials for various purposes (Rushing, Pryma, & Langleben, 2012), often without proof of meeting *Daubert*'s reliability standard (Moriarty, 2008).

Due to constitutional rights, statutory enactments, and concerns over wrongful convictions, criminal defendants may be able to introduce fMRI lie detection testimony without meeting either the *Frye* or the *Daubert* standards in two ways: either in the penalty phases of capital cases, where defendants have a constitutional right to present mitigating evidence (*Smith v. Spisak*, 2010); or to support a claim of postconviction innocence where there is other, newly discovered evidence.

In capital cases, courts frequently permit defendants to introduce a variety of evidence (including neuroscience) to prove brain damage or mental impairment without stringent proof of reliability (Moriarty, 2008). For example, courts have admitted PET and SPECT scans during the penalty phase of capital cases to establish the defendant's mental impairment, even when such evidence may not rise to the level of evidentiary reliability. The U.S. Supreme Court has consistently affirmed constitutional protections for defendants to introduce mitigating evidence in penalty hearings AQ: 26 (*McKoy v. North Carolina*, 1990). "[S]tates cannot limit the sentencer's consideration of any relevant circumstances that could cause it to decline to impose the [death] penalty" (*McCleskey v.* AQ:27,28 *Kemp*, 1987, p. 281). More particularly, the juror may "not be precluded from considering, as a mitigating factor, any aspect of a defendant's character or record and any of the circumstances of the offense that the defendant proffers as a basis for a sentence less AQ:29,30 than death" (*Penry v. Lynaugh*, 1989, p. 317). A defendant may be able to make a compelling case that fMRI lie detection will meet this foregoing standard.

Although the FRE do not apply in sentencing proceedings, some courts have required proof of the reliability of evidence admitted in sentencing (*United States v. Smith*, 2010). This reliability requirement has been mentioned in capital case penalty hearings, upholding the exclusion of polygraph evidence (*United States v. Fulks*, 2006). Because the only cases addressing fMRI evidence of lie detection have found it both unreliable and not generally accepted, courts may not be receptive to the testimony, even in the penalty phase. However, in light of the often lax standards for evidentiary reliability in the penalty phase, the frequent admission of nuclear medicine evidence in these hearings, and the strong constitutional support for defendants' right to introduce mitigating evidence, it is possible that fMRI lie detection evidence will gain a foothold in the courtroom in this manner. For example, a court might permit fMRI evidence that the defendant is being truthful when he expresses remorse about a crime or denies remembering a crime because he was intoxicated. It is thus conceivable that either a trial court will permit such evidence or that an appellate court will find an abuse of discretion where a trial court refused to allow such evidence.

One court has already admitted fMRI evidence relevant to another concern in a penalty hearing. In a 2009 death-penalty case AQ: 31 in Illinois, the defense introduced expert testimony during the penalty phase that the defendant, Brian Dugan, suffered from psychopathy that impaired his ability to control his impulse to kill (Hughes, 2010). The trial court permitted the expert to discuss the fMRI scans taken of Dugan's brain as additional proof of the defendant's psychopathy and to establish that Dugan's psychopathy should make him less culpable. The trial court allowed the expert to explain the scans and to use diagrams of the brain, but disallowed use of the actual fMRI images of Dugan's brain activity. Despite the admission of such expert testimony, Dugan received the death penalty. However, a signed verdict form was discovered after the sentencing, indicating that the jury may have intended to render a verdict of life (Barnum and St. Clair, 2009). If the jury did originally decide not to impose the death penalty, it suggests the testimony was influential. However, Dugan's appeal on this issue was dropped when Illinois abolished the death penalty (Barnum, 2009), so the issue remains unresolved.

fMRI lie detection evidence also has the potential to be admitted posttrial in a compelling case of claimed innocence. In *Harrington v. State* (2003), a trial court permitted testimony from an expert who testified about "brain fingerprinting"—a form of EEG that claims to be able to determine whether a person recognizes a word or image. Although brain fingerprinting has been roundly criticized (Rosenfeld, 2005), the trial court in that case heard testimony from Dr. Farwell, who testified that defendant's brain waves were consistent with his claims of innocence and his alibi. The trial court ultimately denied Harrington's claims, believing them time barred, but the Supreme Court of Iowa reversed, holding that the defendant was entitled to a new trial. On reviewing the record *de novo* and considering all the circumstances, the court's confidence in the soundness of the defendant's conviction was "significantly weakened." Although the Supreme Court of Iowa mentioned Dr. Farwell's testimony in a footnote, it neither commented on the appropriateness of its admission nor relied on it in its decision. It is difficult, however, for defendants to get a new trial after conviction and appeal (Griffin, 2009), and other defendants who sought to hire Farwell met with judicial resistance (Moriarty, 2008). However, another court in a similar circumstance might be more impressed by fMRI evidence, which is based on far more reliable science than the brain fingerprinting (Rosenfeld, 2005; Schauer, 2010b).

The Current State of Scientific Concerns: What Needs to Be Done

Irrespective of which party seeks to introduce the testimony or in what circumstances the proposed testimony is presented, the published indicia of accuracy and reliability of fMRI lie detection are not sufficient for the courtroom. The problem posed and answered here is how to bridge the gap between the basic studies done to date and a requisite standard of evidentiary reliability.

Under certain, controlled laboratory conditions, endorsed lie and truth were distinguished in individual subjects with 76−90% accuracy (Ganis, Rosenfeld, Meixner, Kievit, & Schendan, 2011; Langleben et al., 2005). These findings have been moderated by two recent studies. In the first, Kozel et al. (2009) used a sequence of two deception paradigms generating tasks that involved denying mock crimes. The first mock crime was the scenario from Kozel et al.'s earlier study (2005), in which participants pretended to steal a watch or a ring. fMRI was able to correctly classify 25 of 36 (69%) participants. Those participants whose lies were correctly identified then committed another mock crime and were compared with a control group that did not commit any mock crimes. All participants correctly identified on the first mock crime task were also identified on the second task. However, of the control group, only 5 of 15 were correctly identified, yielding 100% sensitivity but only 33% specificity. Such data could be used to argue that an fMRI test that does not detect deception is highly likely to be correct (Kozel et al., 2009).

Another study using a within-subject design and a sophisticated nonparametric analysis (Davatzikos et al., 2005) reported a classification accuracy of 100% (Ganis et al., 2011), although researchers found it to be reduced to 33% when participants used hand movements as countermeasure. These diverse scenarios, fMRI designs, and data analysis approaches do not allow a direct comparison or an estimate of the overall error rates of the technology. Moreover, they raise the question whether overall error rates are a meaningful variable or whether error rates for each testing scenario need to be evaluated separately.

What is important is that the group differences between lie and truth consistently involve the lateral and inferior prefrontal and posterior parietal cortices, and appear unaffected by gender, handedness, and language. Although this is a fairly advanced state of basic science for a topic in behavioral fMRI research, legitimate forensic use requires substantially more validation. The major issues are validation in ecologically valid situations, where (a) stakes are higher, (b) the more significant potential confounds (subject's age, medical condition, culture) are accounted for (Bizzi et al., 2009; Langleben, 2008; Simpson, 2008; Spence, 2008), and (c) the effects of motor and cognitive countermeasures are evaluated in a deliberate fashion.

Finally, although the inherent accuracy of lie detection within an individual subject is a prerequisite for further translational research, understanding the error rate of a test is not complete until its positive and negative predictive power are also known. The accuracy of discrimination between two conditions within subjects is not equivalent to the probability of detecting liars in a cohort containing liars and truth-tellers, with truth-tellers being a majority. Though studies have begun to address these gaps (Abe et al., 2009; Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Ganis et al,. 2011; Kaylor-Hughes et al., 2011; Kozel et al., 2009; Mildner, Zysset, Trampel, Driesel, & Moller, 2005; Nuñez, Casey, Egner, Hare, & Hirsch, 2005), comprehensive answers to the translational questions require a more robust effort.

Several technical aspects of cognitive fMRI experiments have direct forensic relevance and raise additional questions that researchers might develop. First, BOLD fMRI, used in all fMRI studies of deception so far, is one of many fMRI techniques, and fMRI itself is one of many approaches available on most high-field MRI scanners. Conceivably, other fMRI approaches could be superior to BOLD in lie detection. Second, as we mentioned earlier, BOLD fMRI provides a relative rather than an absolute measure of local brain activity. Consequently, lie detection using BOLD fMRI involves interpretation of the BOLD fMRI signal differences between a test and comparison questions; the questions used are thus critical to the result. Third, it is unclear whether there is a brain fMRI pattern specific to deception, and at least some of the studies indicate that the pattern of deception is specific to the experimental paradigm used to generate it. Though the left prefrontal cortex is a leading candidate for a region specifically activated during deception (Spence & Kaylor-Hughes, 2008; Spence, Kaylor-Hughes, Brook, Lankappa, & Wilkinson, 2008), until these data are clinically validated, we cannot assume that fMRI patterns and error rates will generalize between deception tasks with different sequence and content of target or comparison questions. fMRI discrimination between lie and truth is possible without knowing whether there is a deception-specific fMRI pattern, as long as the difference between lie and truth in a specific questioning format (i.e., the CIT) is known and reliable. This reliance of fMRI-based lie detection on discrimination between two behavioral conditions (lie and a known truth or other baseline) generated by a preset question format allows the translational studies of clinical relevance to proceed without waiting for the outcomes of the search for the "lie center" in the brain. This question is part of the debate about localized versus distributed functions in the brain that dates back to the 19th century, and it may continue well after the determination of the utility and scope of the potential use of fMRI for lie detection is concluded. As an analogy, we use antidepressant drugs extensively, without knowing their exact mechanisms of action.

Though the cognitive neuroscience investigation of deception is clinically relevant in the long run, the critical question of error rates and other translational questions described earlier can and should be answered independent of the basic research on the mechanisms, because they will determine the level of public interest in the entire field of fMRI based lie detection. Similarly, though the interaction among memory, emotion, and deception is important academically (Phelps, 2009), and for the comprehensive understanding of the countermeasures to lie detection, the translational studies can proceed ahead of or simultaneously with basic research. It is also likely that many of the basic science questions on the mechanisms of deception could be incorporated into clinical trials with no added costs.

The Current State of Legal Concerns: What Issues Remain Unresolved?

Scholars continue to discuss factors affecting admissibility related to both scientific and legal considerations. For example, Shen and Jones (2011) have focused on the design of the tasks, the ecological and external validity of the conditions, and concerns about statistical methods and group-data averaging implications. Other voiced concerns include data interpretations and the problems of ecological validity (Kanwisher, 2009), as well as the various juridical concerns that neuroscience lie detection—like other forms of lie detection—pose for courts (Rakoff, 2009; Imwinkelried, 2011). AQ: 32

Although we believe the concerns raised in *Semrau* address the primary considerations related to legal reliability, it is also worth noting that complications arise from discrepancies in the meaning of crucial terms such as validity and reliability between law and science. For example, in medicine and biostatistics, the term *validity* refers to the relevance of the test, that is, whether the test actually measures what it purports to measure. For example, to determine whether fMRI lie detection is a valid test of deception, one would ask whether the brain activation detected by fMRI during a deception task is related to deception. The term *reliability* refers to reproducibility of the test results when the test is repeated. With fMRI lie detection, this would mean that the same regions of the brain repeatedly show activation when presented with the same question within a single session and across several different sessions.

Courts and litigants, however, do not assign the same meaning to reliability or use it with the scientific level of precision. For example, when lawyers argue about the "reliability" of expert evidence, they debate whether the testimony is sufficiently "trustworthy" to constitute appropriate courtroom evidence; they rarely are referring to its reproducibility. A colorful example of the law's interpretation of reliability is found in Justice Scalia's concurrence in *Kumho Tire Co., Inc. v. Carmichael* (1999), where he noted that the court has the discretion "to choose among *reasonable* means of excluding expertise that is *fausse* and science that is junky" (p. AQ: 33 159).

In *Daubert v. Merrell Dow Pharmaceuticals, Inc*. (1993), the U.S. Supreme Court commented that "to qualify as 'scientific knowledge,' an inference or assertion must be derived by the scientific method. Proposed testimony must be supported by appropriate validation—that is, 'good grounds,' based on what is known. In short, the requirement that an expert's testimony pertain to 'scientific knowledge' establishes a standard of evidentiary AQ: 34 reliability" (p. 590). In footnote 9 following the above quotation, the Court explained its understanding of the distinction between reliability and validity:

> We note that scientists typically distinguish between "validity" (does the principle support what it purports to show?) and "reliability" (does application of the principle produce consistent results?). . . . Although "the difference between accuracy, validity, and reliability may be such that each is distinct from the other by no more than a hen's kick," . . . our reference here is to *evidentiary* reliability—that is, trustworthiness. . . . In a case involving scientific evidence, *evidentiary reliability* AQ:35,36 will be based upon *scientific validity.* (p. 590, n.9)

Thus the Court defined "legal reliability" in terms of "scientific validity." Although this muddling of the terms may have been intentional, it is equally probable that the Court was aiming at the *concept* of validity: Does the test actually do what it purports to do? Analyzing this standard in terms of the legal reliability of fMRI lie detection, the question is the same: Does the fMRI test determine whether a given person is or is not lying? The only answer that current data can provide is that, in a controlled laboratory setting, fMRI can identify deceptive responses with 76–90% accuracy. Is that enough for "legal reliability"? We do not think so. Without knowing the predictive power of the test in an ecologically-valid setting, there is no accurate way to respond to *Daubert*'s "known error rate" inquiry. This science is currently in AQ: 37 the area focused on in the *Joiner* court, where the Court remarked that there may be "simply too great an analytic gap between the data and the opinion proffered" (*General Electric Co. v. Joiner*, AQ: 38 1997, p. 146). Until properly controlled trials are done, the science remains in that "analytic gap." But such current concerns about fMRI lie detection are not fatal to the endeavor—rather, the science is in its nascent form and requires time and funding to better define its clinical potential. Similar critiques were leveled at early studies conducted on eyewitness identification, which, after much continued research, now qualifies as scientifically reliable AQ: 39 evidence (Cutler & Wells, 2009; Leippe, 1995). Despite the need for a good method to detect deception, we do not have one, and "the research should vigorously explore alternatives to the polygraph, AQ: 40 including functional brain imaging" (Raichle, 2009, p. 6).

A major concern with fMRI lie detection is the looming problem that subsequent studies will prove the early studies wrong—a possible outcome in all developing research. The danger of admitting scientific evidence before it is proven to be sufficiently reliable and valid is by now well known. For example, Garrett and Neufeld (2009) examined the trial transcripts of 137 exonerated defendants and concluded that approximately 60% of those trials included flawed science. The NRC Report (2009) concluded that "no forensic method other than nuclear DNA analysis has been rigorously shown to have the capacity to consistently and with a high degree of certainty support conclusions about . . . 'matching' of an unknown item of evidence to a specific known source" (p. 87). These forms of forensic evidence include fingerprints, tool AQ: 41 marks, handwriting, bite marks, and hair comparison—the often critical evidence in criminal trials.

The NRC Report (2009) finds that the interpretation of forensic evidence is not always based on scientific studies to determine its validity. There is no body of research on the limits and measures of performance, the analysts have differing skill levels, there are potential bias concerns among those performing the analysis, and no rigorous protocols to guide the subjective interpretations. The NRC Report goes so far as to say "[t]he law's greatest dilemma in its heavy reliance on forensic evidence, however, concerns the question of whether—and to what extent—there is any *science* in any given 'forensic science' discipline" (p. 87). Despite the scath- AQ: 42 ing critique of such forensic science, it continues to flow into the courtroom (Moriarty, 2010). For example, recent decisions have upheld the admission of both fingerprint and handwriting comparison, despite recognition of the NCR Report's criticisms (fingerprint comparison: *United States v. Love*, 2011; handwriting: *Pettus v. United States*, 2012).

Even at this early stage, the fMRI lie detection research is far better grounded than much of what passes for forensic science, as the more than two dozen peer-reviewed articles on the subject have established. Nonetheless, we do not believe it is ready for the courtroom. Such potentially powerful testimony as fMRI lie detection should not be admissible without better proof of validity and reliability. The courts are now grappling with forensic science that has been admitted without adequate proof of reliability; we should not repeat this error with fMRI lie detection.

The Policy Analysis

We do believe that fMRI offers a theoretical possibility of improvement over current means of credibility assessment, and could satisfy the yet unmet needs of the legal, defense, and law enforcement communities (NRC Report, 2009). Objective means for detecting deception have a high potential for social benefit. Moreover, fMRI studies of deception have provided important scientific insights into the role of deception in cognition (Greene & Paxton, 2009; Langleben et al., 2002) that are relevant to such diverse topics as morality, drug addiction, and treatment nonadherence in chronic medical illness. Thus, the topic is well worth pursuing with both translational and basic research. Therefore, it is in the public interest to guide the development of fMRI lie detection technology, rather than leave it to other stakeholders, such as for-profit companies.

Though companies offering commercial MRI veracity testing seem to promise more than they can deliver, we do not believe that new legislation is needed to regulate their activity or the admissibility of their data as evidence. First, the size and scope of these companies is exceedingly small. Second, despite substantial problems of reliability and jurisprudential concerns about the polygraph, there has been no major movement to legislatively ban its use in all circumstances, except for nongovernment preemployment testing (Office of Technology Assessment, 1990) and other limited categories, and certainly not to prohibit it as a category of evidence. Third, there has been no apparent movement to enact an FRE provision similar to Military Rule of Evidence (MRE) 707, which bans polygraph evidence, despite the U.S. Supreme Court upholding the constitutionality of MRE 707 (*United States v. Scheffer*, 1998). Finally, there is likely little political interest in championing legislative prohibition about fMRI, given the current state of political affairs and the more critical public interest concerning the substantial shortcomings of forensic science currently in use. Rather than focus on regulation, we propose to use science to pull fMRI lie detection out of the limbo. Specifically, practical legal analysis and comprehensive translational experimental data are needed to resolve the remaining questions of fMRI veracity testing.

The most important missing piece in the puzzle is *Daubert*'s "known error rate" standard. Determining the error rates for fMRI-based lie detection requires validation of the method in settings convincingly approximating the real-life situations in which legally significant deception takes place, in terms of the risk−benefit ratio, relevant demographics, and the prevalence of the behavior in question. AQ: 43

Clinical validation of a test is an expensive enterprise, often performed by commercial interests. Under the medical model of drug and device development, controlled clinical trials are required to determine whether the device is efficacious and superior to existing alternatives, and to determine the error rates in the target populations. Applied to fMRI lie detection, such trials would include testing the technology in key target populations and age groups under deception scenarios with various levels of risk and benefit. This implies that some of the trials would have to hold the deception scenario constant while testing the effect of a demographic variable on the outcomes, while others would have to hold the demographic constant and manipulate the experimental scenario or task. The relatively large number of variables is likely to require the large overall number of participants, though the number required for each study could be relatively small (e.g., 30–100). Continuing the parallel to medical test development, the incidence of spontaneous deception in the target populations is variable and rather low. Baldessarini et al. (1983) elaborated on the potential clinical validation of the Dexamethasone Suppression Test (DST). In a research setting, DST had 70% sensitivity and 95% specificity for diagnosing depression. In Baldessarini et al.'s example, the positive predictive value (PPV) was 93% in the research sample that had a 50% prevalence of the disease (100 patients with depression and 100 healthy controls). In a specialty clinic, where the prevalence of depression was 10%, the PPV of the test declined to 63%, and in a primary care setting, with sample of 1,000 and disease prevalence of 1%, the PPV became a dismal 12% (Baldessarini et al., 1983).

We draw three conclusions from this illustration. First, screening settings are more demanding on test accuracy, and it is unlikely that fMRI or any other lie detector, including the polygraph, will ever reach the positive predictive power sufficient for screening for deception among large groups of mostly innocents. Second, fMRI-based technology may be useful in the forensic settings where the prevalence of deception is much higher than in the general population. Third, a series of properly powered and controlled prospective studies (i.e., clinical trials) would be required to confirm or disprove this hypothesis. Such studies would be adequately powered to include a few target participants (liars) mixed into a proportionally large number of honest participants. This would permit meaningful calculations of the error rates, including within-subject accuracy and predictive values. Despite the ethical challenges such trials may pose, forensic functional imaging studies are not inconceivable in both normal and pathological populations (Fullam, McKie, & Dolan, 2009; Hakun et al., 2009; Kozel & Trivedi, 2008; Yang et al., 2007). Another way of estimating new technology's efficacy is a head-to-head comparison between fMRI and the polygraph. Finally, mathematical modeling could help extrapolate findings. Such studies would involve hundreds of participants and would cost millions of dollars but far less than the average Phase III clinical trial. Though a recommendation for more research may seem too general, guiding fMRI lie detection research toward socially beneficial and conclusive findings is unlikely to occur without targeted policy.

In clinical development terms, the fMRI lie detector is stuck between Phase I and Phase II clinical trials, with the commercial start-ups lacking the capacity to proceed to Phase III—a common situation with compounds or devices of unclear commercial value. For devices with clear public health interest, such as medications for drug addiction, the U.S. National Institutes of Health (NIH) have often bridged the funding gap. Despite a pivotal role of deception in a range of personality disorders, drug and alcohol abuse, and treatment nonadherence, so far NIH has not recognized deception as a health issue. U.S. defense and intelligence agencies have funded research in this area, but results have been slow to appear in the scientific literature (Dedman, 2009; Stern, 2003) and may be subject to nonscientific bias similar to those that afflicted the U.S. Department of Defense−sponsored polygraph research. A $5 million congressional earmark in the 2004 and 2005 defense budgets funded the Center for Advanced Technologies for Deception Detection (CATDD) at the University of South Carolina at Columbia (Hickman, 2005). At the time of this writing, we were unable to identify peer-reviewed publications on lie detection from CATDD. The MacArthur Foundation's $10 million Law and Neuroscience Project (2010) has produced some important basic data on lie detection (Greene & Paxton, 2009) with ethical and legal analysis, but has not addressed the translational questions (Gazzaniga, 2008). Thus, no group has been able to spur the program of translational research outlined above, while the clinical nature and relatively large scope puts such a project outside of the usual purview of the National Science Foundation. AQ: 44

Conclusion

We believe that, at the present stage of development, the most important policy intervention in the field of brain-based lie detectors is a public funding initiative leading to a peer-reviewed translational research program with a special emphasis on a series of multicenter clinical trials to determine the error rates of the technique, the sensitivity to countermeasures, the effect of high benefit-to-risk ratios, the relative accuracy compared to the polygraph, and the effects of age, gender, common pharmacological agents, and cognitive status. The specificity of any given pattern of brain activity to deception is likely to be addressed as a byproduct of the studies described above. Considering the controversial nature of the topic and the potential societal impact of this technology, a collaboration of several agencies may be required to create a funding mechanism that could impartially assess and guide the development of forensic fMRI technology.

References

Abe, N., Fujii, T., Hirayama, K., Takeda, A., Hosokai, Y., Ishioka, T., . . . Mori, E. (2009). Do parkinsonian patients have trouble telling lies? The neurobiological basis of deceptive behaviour. *Brain, 132,* 1386–1395. doi:10.1093/brain/awp052

Aguirre, G. K., & D'Esposito, M. (1999). Experimental design for brain fMRI. In C. T. W. Moonen & P. A. Bandettini (Eds.), *Functional MRI* (pp. 369–380). New York, NY: Springer.

Baldessarini, R. J., Finklestein, S., & Arana, G. W. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry, 40,* 569–573. doi:10.1001/archpsyc.1983.01790050095011

Barnum, A. (2009, March 11). Dugan asks to drop appeal in Nicarico case. *Chicago Tribune.* Retrieved from http://articles.chicagotribune.com/2011-03-11/news/ct-met-dugan-appeal-0312-20110311_1_brian-dugan-jeanine-nicarico-nicarico-family

Barnum, A., & St. Clair, S. (2009, December). Dugan juror who switched at last minute stands by death sentence. *Chicago Tribune.* Retrieved from http://articles.chicagotribune.com/2009-12-03/news/0912020807_1_death-sentence-two-jurors-death-penalty

Ben-Shakhar, G. (1991). Clinical judgment and decision-making in CQT-polygraphy. A comparison with other pseudoscientific applications in psychology. *Integrative Physiological & Behavioral Science, 26,* 232–240. doi:10.1007/BF02912515

Ben-Shakhar, G. (2001). *A critical review of the Controlled Question Test (CQT).* In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 103–127). San Diego, CA: Academic Press.

Ben-Shakhar, G., Bar-Hillel, M., & Kremnitzer, M. (2002). Trial by polygraph: Reconsidering the use of the guilty knowledge technique in court. *Law and Human Behavior, 26,* 527–541. doi:10.1023/A:1020204005730

Berger, H. (1929). Über das Elektroenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten, 87,* 527–570. doi:10.1007/BF01797193

Bizzi, E., Hyman, S., Raichle, M., Kanwisher, N., Phelps, E., Morse, S., . . . Greely, H. T. (Eds.). (2009). *Using imaging to identify deceit: Scientific and ethical questions.* Cambridge, MA: American Academy of Arts & Sciences.

Canli, T., Brandon, S., Casebeer, W., Crowley, P. J., Du Rousseau, D., Greely, H. T., & Pascual-Leone, A. (2007). Neuroethics and national security. *American Journal of Bioethics, 7,* 3–13. doi:10.1080/15265160701290249

Crewson, P. E. (2001). *Comparative analysis of polygraph with other screening and diagnostic tools* (Report No. DODPI00P0030). Ashburn, VA: Research Support Service.

Cutler, B. L., & Wells, G. L. (2009). Psychological science in the courtroom: Consensus and controversy. In J. L. Skeem & K. S. Douglas (Eds.), *Expert testimony regarding eyewitness identification* (pp. 100–123). New York, NY: Guilford Press. doi:10.1093/acprof:oso/9780195331974.001.0001

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 113 S. Ct. 2786, 125 L. Ed. 2d 469 (1993).

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., . . . Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage, 28,* 663–668. doi:10.1016/j.neuroimage.2005.08.009

Dedman, B. (2008). New anti-terror weapon: Hand-held lie detector. U.S. troops in Afghanistan first to get new device; 'red' means you're lying. Retrieved from http://www.msnbc.msn.com/id/23926278

Donaldson, D. I., & Buckner, R. L. (2001). Effective paradigm design. In M. P. M. Matthews, P. Jezzard, & A. C. Evans (Eds.), *Functional MRI: An introduction to methods* (pp. 176–195). New York, NY: Oxford University Press.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46,* 913–920. doi:10.1037/0003-066X.46.9.913

Faigman, D. L., Kaye, D. H., Saks, M. J., & Sanders, J. (2010). *Modern scientific evidence: The law and science of expert testimony* (§§ 40.1, 17:10). Eagan, MN: Thomson West. AQ: 47

Fisher, G. (1997). The jury's rise as lie detector. *Yale Law Journal, 107,* 575–713. doi:10.2307/797252

Fox, D. (2009). The right to silence as protecting mental control. *Akron Law Review, 42,* 763–801.

Frye v. United States, 293 F. 1013 (D. C. Cir. 1923).

Fullam, R. S., McKie, S., & Dolan, M. C. (2009). Psychopathic traits and deception: Functional magnetic resonance imaging study. *British Journal of Psychiatry, 194,* 229–235. doi:10.1192/bjp.bp.108.053199

Gallini, B. R. (2010). Police science in the interrogation room: Seventy years of pseudo-psychological interrogation methods to obtain inadmissible confessions. *Hastings Law Journal, 61,* 529–580.

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An FMRI investigation. *Cerebral Cortex, 13,* 830–836. doi:10.1093/cercor/13.8.830

Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *NeuroImage, 55,* 312–319. doi:10.1016/j.neuroimage.2010.11.025

Garrett, B. L., & Neufeld, P. J. (2009). Invalid forensic science testimony and wrongful conviction. *Virginia Law Review, 95,* 1–76.

Gazzaniga, M. S. (2008). The law and neuroscience. *Neuron, 60,* 412–415. doi:10.1016/j.neuron.2008.10.022

General Electric Co. v. Joiner, 522 U.S. 136, 118 S. Ct. 512, 139 L. Ed. 2d 508 (1997).

Gjedde, A. (2001). Brain energy metabolism and the physiological basis of the haemodynamic response. In P. Jezzard, P. M. Matthews, & S. Smith (Eds.), *Functional MRI: An introduction to methods* (pp. 37–66). New York, NY: Oxford University Press.

Greely, H. T. (2006). The social effects of advances in neuroscience: Legal problems, legal perspective. In J. Illes (Ed.), *Neuroethics: Defining the issues in theory, practice and policy* (pp. 245–263). New York, NY: Oxford University Press.

Greely, H. T., & Illes, J. (2007). Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law & Medicine, 33,* 377–431.

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 106,* 12506–12511. doi:10.1073/pnas.0900152106

Griffin, L. (2009). Correcting injustice: How the United Kingdom and the United States review claims of innocence. *Toledo Law Review, 41,* 107–169.

Grubin, D. (2010). The polygraph and forensic psychiatry. *Journal of the American Academy of Psychiatry and the Law, 38,* 446–451.

Hakun, J. G., Ruparel, K., Seelig, D., Busch, E., Loughead, J. W., Gur, R. C., & Langleben, D. D. (2009). Towards clinical trials of lie detection with fMRI. *Social Neuroscience, 4,* 518–527. doi:10.1080/17470910802188370

Halber, D. (2007, February 14). Scientists swear a good lie detector is hard to find. *TechTalk, 51*(17), 7. Retrieved from http://web.mit.edu/newsoffice/2007/techtalk51–17.pdf

Halliburton, C. (2009). How privacy killed Katz: A tale of cognitive freedom and the property of personhood as fourth amendment norm. *Akron Law Review, 42,* 803–884.

Happel, M. D. (2005). Neuroscience and the detection of deception. *Review of Policy Research, 22,* 667–685. doi:10.1111/j.1541-1338.2005.00166.x

Harrington v. State, 659 N. W. 2d 509 (Iowa 2003).

Henig, R. M. (2006, February 5). The new science of lying. Brain mapping, E.R.P. waves and microexpression readings could make deception harder. But would that make us better? *New York Times Magazine, 47.*

AQ: 50

Hickman, W. (2005). Senate approves $100 million for South Carolina defense projects. Retrieved from http://lgraham.senate.gov/public

Hughes, V. (2010). Science in court: Head case. *Nature, 464,* 340–342. doi:10.1038/464340a

Hyman, S. E. (2010). Emerging neurotechnologies for lie-detection: Where are we now? An appraisal of Wolpe, Foster and Langleben's "Emerging neurotechnologies for lie-detection: Promise and perils" five years later. *American Journal of Bioethics, 10,* 49–50. doi:10.1080/15265161.2010.527263

Imwinkelried, E. J. (2011). Serendipitous timing: The coincidental emergence of the new brain science and the advent of an epistemological approach to determining the admissibility of expert testimony. *Mercer Law Review, 62,* 959–987.

Kanwisher, N. (2009). The use of fMRI in lie detection: What has been shown and what has not. In E. Bizzi, S. Hyman, M. Raichle, N. Kanwisher, E. Phelps, S. Morse, . . . H. T. (Eds.), *Using imaging to identify deceit: Scientific and ethical questions* (pp. 7–13). Cambridge, MA: American Academy of Arts & Sciences.

Kaye, D. H., Bernstein, D. E., & Mnookin, J. L. (2012). §2.4 Expert testimony on credibility: General rule. In *The new Wigmore: A treatise on evidence − Expert evidence* (2nd. ed., pp. 1–4). Frederick, MD; Aspen Publishers.

Kaylor-Hughes, C. J., Lankappa, S. T., Fung, R., Hope-Urwin, A. E., Wilkinson, I. D., & Spence, S. A. (2011). The functional anatomical distinction between truth telling and deception is preserved among people with schizophrenia. *Criminal Behaviour and Mental Health, 21,* 8–20. doi:10.1002/cbm.785

Kozel, F. A., Johnson, K. A., Grenesko, E. L., Laken, S. J., Kose, S., Lu, X., . . . George, M. S. (2009). Functional MRI detection of deception after committing a mock sabotage crime. *Journal of Forensic Sciences, 54,* 220–231. doi:10.1111/j.1556-4029.2008.00927.x

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry, 58,* 605–613. doi:10.1016/j.biopsych.2005.07.040

Kozel, F. A., & Trivedi, M. H. (2008). Developing a neuropsychiatric functional brain imaging test. *Neurocase, 14,* 54–58. doi:10.1080/13554790701881731

Kumho Tire Co., Ltd. v. Carmichael, 526 U.S. 137, 119 S. Ct. 1167, 143 L. Ed. 2d 238 (1999).

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., . . . Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 89,* 5675–5679. doi:10.1073/pnas.89.12.5675

Langleben, D. D. (2008). Detection of deception with fMRI: Are we there yet? *Legal and Criminological Psychology, 13,* 1–9. doi:10.1348/135532507X251641

Langleben, D. D., Loughead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping, 26,* 262–272. doi:10.1002/hbm.20191

Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., . . . Childress, A. R. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *NeuroImage, 15,* 727–732. doi:10.1006/nimg.2001.1003

Langleben, D. D., Willard, D., Moriarty, J. C. (2012). Detecting deception. In J. R. Simpson (Ed.), *Neuroimaging in forensic psychiatry: From the clinic to the courtroom* (pp. 217–236). Chichester, England: John Wiley.

AQ: 51

Lauterbur, P. C. (1973). Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature, 242,* 190–191. doi:10.1038/242190a0

Lee, T. M., Liu, H. L., Tan, L. H., Chan, C. C., Mahankali, S., Feng, C. M., . . . Gao, J. H. (2002). Lie detection by functional magnetic resonance imaging. *Human Brain Mapping, 15,* 157–164. doi:10.1002/hbm.10020

Leippe, M. R. (1995). The case for expert testimony about eyewitness memory. *Psychology, Public Policy, and Law, 1,* 909–959.

Lykken, D. T. (1991). Why (some) Americans believe in the lie detector while others believe in the guilty knowledge test. *Integrative Physiological & Behavioral Science, 26,* 214–222. doi:10.1007/BF02912513

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107,* 343–352.

McCleskey v. Kemp, 481 U.S. 279, 281 (1987).

McKoy v. North Carolina, 494 U.S. 433 (1990). AQ: 52

Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professional and lawyers* (3rd ed.). New York, NY: Guilford Press.

Mildner, T., Zysset, S., Trampel, R., Driesel, W., & Moller, H. E. (2005). Towards quantification of blood-flow changes during cognitive task activation using perfusion-based fMRI. *NeuroImage, 27,* 919–926. doi:10.1016/j.neuroimage.2005.04.040

Miller, G. R., & Stiff, J. B. (1993). Investigating deceptive communication. In *Deceptive communication* (Vol. 14, pp. 32–39). Newbury Park, CA: Sage.

Monahan, J., Walker, L., & Mitchell, G. (2008). Contextual evidence of gender discrimination: The ascendance of "social frameworks." *Virginia Law Review, 94,* 1715–1748.

Moriarty, J. C. (2008). Flickering admissibility: Neuroimaging evidence in the U.S. courts. *Behavioral Sciences & the Law, 26,* 29–49. doi:10.1002/bsl.795

Moriarty, J. C. (2009). Visions of deception: Neuroimages and the search for truth. *Akron Law Review, 42,* 739–762.

Moriarty, J. C. (2010). Will history be servitude? The NAS Report on forensic science and the role of the judiciary. *Utah Law Review, 2010,* 299–326.

Myers, J. E. B. (2010). Expert testimony in child sexual abuse litigation: Consensus and confusion. *University of California Davis Journal of Juvenile Law & Policy, 14,* 1–57.

Nakayama, M. (2001). Practical of the concealed information test for criminal investigations in Japan. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 49–86). San Diego, CA: Academic Press.

Nance, D. A., & Morris, S. B. (2002). An empirical assessment of presentation formats with a relatively large and quantifiable random match probability. *Jurimetics Journal, 42,* 403–445.

Nance, D. A., & Morris, S. B. (2005). Juror understanding of DNA evidence: An empirical assessment of presentation formats for trace evidence with a relatively small random-match probability. *Journal of Legal Studies, 34,* 395–444. doi:10.1086/428020

National Research Council of the National Academies. (2009). *Strengthening forensic science in the United States: A path forward.* Washington, D. C.: National Academies Press.

Nicolas, P. (2004). De novo review in deferential robes?: A deconstruction of the standard of review of evidentiary errors in the federal system. *Syracuse Law Review, 54,* 531–533.

Nuñez, J. M., Casey, B. J., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *NeuroImage, 25,* 267–277. doi:10.1016/j.neuroimage.2004.10.041

Office of Technology Assessment. (1990). *The use of integrity tests for pre-employment screening* (No. Report No. OTA-SET-442). Washington, DC: Author.

Oullier, O. (2012). Clear up this fuzzy thinking on brain scans. *Nature, 483,* 7. doi:10.1038/483007a

Owen, A. M., Epstein, R., & Johnsrude, I. S. (2001). fMRI: Applications in cognitive neuroscience. In P. Jezzard, P. M. Matthews, & S. Smith (Eds.), *Functional MRI: An introduction to methods* (pp. 311–327). New York, NY: Oxford University Press.

Pardo, M. S. (2006). Neuroscience evidence, legal culture, and criminal procedure. *American Journal of Criminal Law, 33,* 301–335.

AQ: 53

Penry v. Lynaugh, 492 U.S. 302, 317 (1989).

Pettus v. United States, _A. 3d._, 2012 WL 399997 (D. C. 2012).

Phelps, E. A. (2009). Lying outside the laboratory: The impact of imagery and emotion on the neural circuitry of lie detection. In E. Bizzi, S. Hyman, M. Raichle, N. Kanwisher, E. Phelps, S. Morse, . . . H. T. (Eds.), *Using imaging to identify deceit: Scientific and ethical questions* (pp. 14–22). Cambridge, MA: American Academy of Arts & Sciences.

Poulin, A. B. (2007). Credibility: A fair subject for expert testimony? *Florida Law Review, 59,* 991–1069.

Raichle, M. E. (2009). An introduction to functional brain imaging in the context of lie detection. In E. Bizzi, S. Hyman, M. Raichle, N. Kanwisher, E. Phelps, S. Morse, . . . H. T. (Eds.), *Using imaging to identify deceit: Scientific and ethical questions* (pp. 3–6). Cambridge, MA: American Academy of Arts & Sciences.

Rakoff, J. S. (2009). Lie detection in the courts: The vain search for the magic bullet. In E. Bizzi, S. Hyman, M. Raichle, N. Kanwisher, E. Phelps, S. Morse, . . . H. T. (Eds.), *Using imaging to identify deceit: Scientific and ethical questions* (pp. 40–45). Cambridge, MA: American Academy of Arts & Sciences.

Raskin, D. C., & Honts, C. R. (2001). The comparison question test. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 14–15). San Diego, CA: Academic Press.

Risinger, D. M. (2000). Defining the “task at hand”: Non-science forensic science after Kumho Tire Co. v. Carmichael. *Washington & Lee Law Review, 57,* 767–800.

Rosenfeld, J. P. (2005). “Brain fingerprinting”: A critical analysis. *The Scientific Review of Mental Health Practice, 4*(1), 20–37.

Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience, 42,* 157–161. doi:10.3109/00207458808985770

Rushing, S. E., Pryma, D. A., & Langleben, D. D. (2012). PET and SPECT. In J. R. Simpson (Ed.), *Neuroimaging in forensic psychiatry: From the clinic to the courtroom* (pp. 3–25). Chichester, England: John Wiley. doi:10.1002/9781119968900.ch1

AQ: 54

Schauer, F. (2010a). Can bad science be good evidence? *Cornell Law Review, 95,* 1191–1220.

Schauer, F. (2010b). Neuroscience, lie-detection, and the law: Contrary to the prevailing view, the suitability of brain-based lie-detection for courtroom or forensic use should be determined according to legal and not scientific standards. *Trends in Cognitive Sciences, 14,* 101–103. doi: 10.1016/j.tics.2009.12.004

Schweitzer, N. J., Saks, M. J., Murphy, E. R., Roskies, A. L., Sinnott-Armstrong, W., & Gaudet, L. M. (2011). Neuroimages as evidence in a mens rea defense: No impact. *Psychology, Public Policy, and Law, 17,* 357–393. doi:10.1037/a0023581

Seaman, J. A. (2008). Black boxes. *Emory Law Journal, 58,* 427–488.

Shen, F. X., & Jones, B. D. (2011). Brain Scans as evidence: Truths, proofs, lies, and lessons. *Mercer L.Rev., 62,* 861.

Simpson, J. R. (2008). Functional MRI lie detection: Too good to be true? *Journal of the American Academy of Psychiatry and the Law, 36,* 491–498.

Smith v. Spisak, 130 S. Ct. 676, 175 L. Ed. 2d 595 (2010).

Spence, S. A. (2008). Playing devil’s advocate: The case against fMRI lie detection. [Editorial]. *Legal and Criminological Psychology, 13,* 11–25. doi:10.1348/135532507X251597

Spence, S. A., Farrow, T. F., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. (2001). Behavioural and functional anatomical correlates of deception in humans. *NeuroReport: For Rapid Communication of Neuroscience Research, 12,* 2849–2853. doi:10.1097/00001756-200109170-00019

Spence, S. A., Hope-Urwin, A., Lankappa, S. T., Woodhead, J., Burgess, J. C. L., & Mackay, A. V. (2010). If brain scans really detected deception, who would volunteer to be scanned?, *Journal of Forensic Sciences, 55,* 1352–1355. doi:10.1111/j.1556-4029.2010.01452.x

Spence, S. A., & Kaylor-Hughes, C. J. (2008). Looking for truth and finding lies: The prospects for a nascent neuroimaging of deception. *Neurocase, 14,* 68–81. doi:10.1080/13554790801992776

Spence, S. A., Kaylor-Hughes, C. J., Brook, M. L., Lankappa, S. T., & Wilkinson, I. D. (2008). ‘Munchausen’s syndrome by proxy’ or a ‘miscarriage of justice’? An initial application of functional neuroimaging to the question of guilt versus innocence. *European Psychiatry, 23,* 309–314. doi:10.1016/j.eurpsy.2007.09.001

Stern, P. C. (Ed.). (2003). *The polygraph and lie detection. Report of the National Research Council Committee to Review the Scientific Evidence on the Polygraph.* Washington, DC: The National Academies Press.

Talbot, M. (2007, July 2). Duped: Can brain scans uncover lies? *The New Yorker.* Retrieved from http://www.newyorker.com/reporting/2007/07/02/070702fa_fact_talbot

The Law and Neuroscience Project (2010). Retrieved from http://www.lawneuro.org

Thompson, S. K. (2005). The legality of the use of psychiatric neuroimaging in intelligence interrogation. *Cornell Law Review, 90,* 1601–1637.

Tovino, S. A. (2007). Imaging body structure and mapping brain function: A historical approach. *American Journal of Law and Medicine, 33,* 193–228.

United States v. Charley, 189 F. 3d 1251, 1267 (10th Cir. 1999).

United States v. Fulks, 454 F. 3d 410 (4th Cir. 2006).

United States v. Love, F. 3d, 2011 WL 2173644 (S. D. Cal.).

United States v. Lumpkin, 192 F. 3d 980 (2d Cir. 1999).

United States v. Scheffer, 523 U.S. 303 (1998).

United States v. Semrau, 07–10074 ML/P, 2010 WL 6845092 (W. D. Tenn. June 1, 2010).

United States v. Semrau, 2011 WL 3881885 (corrected appeal brief filed in the Sixth Circuit).

United States v. Sherlin, 67 F. 3d 1208 (6th Cir. 1995).

United States v. Smith, 405 Fed. App’x 795 (4th Cir. 2010).

United States v. Swayze, 378 F. 3d 834, 837 (8th Cir. 2004).

U.S. Senate. (1994). *An assessment of the Aldrich H. Ames espionage case and its implications for U.S. intelligence* (No. ISBN-10 0160460972). Washington, DC: U.S. Congress.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester, England: Wiley.

Walker, L., & Monahan, J. (1987). Social frameworks: A new use of social science in law. *Virginia Law Review, 73,* 559–598. doi:10.2307/1072923

Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20,* 470–477. doi:10.1162/jocn.2008.20040

Wilson v. Corestaff Services, L. P., 900 N. Y. S. 2d 639 (N. Y. Sup. Ct. 2010).

Wolpe, P. R., Foster, K. R., & Langleben, D. D. (2005). Emerging neurotechnologies for lie-detection: Promises and perils. *American Journal of Bioethics, 5,* 39–49. doi:10.1080/15265160590923367

Yang, Y., Raine, A., Narr, K. L., Lencz, T., LaCasse, L., Colletti, P., & Toga, A. W. (2007). Localisation of increased prefrontal white matter in pathological liars. *British Journal of Psychiatry, 190,* 174–175. doi: 10.1192/bjp.bp.106.025056

Received August 6, 2011
Revision received March 20, 2012
Accepted March 27, 2012 ■