



Published in final edited form as:

Psychol Public Policy Law. 2013 May 1; 19(2): 222–234. doi:10.1037/a0028841.

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

Daniel D. Langleben¹ and Jane Campbell Moriarty²

¹Associate Professor, University of Pennsylvania Perelman School of Medicine

²Carol Los Mansmann Chair in Faculty Scholarship, Associate Dean for Faculty Scholarship, and Professor of Law, Duquesne University School of Law

Abstract

Progress in the use of functional magnetic resonance imaging (fMRI) of the brain to evaluate deception and differentiate lying from truth-telling has created anticipation of a breakthrough in the search for technology-based methods of lie detection. In the last few years, litigants have attempted to introduce fMRI lie detection evidence in courts. This article weighs in on the interdisciplinary debate about the admissibility of such evidence, 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.

We believe that the *Daubert's* “known error rate” is the key concept linking the legal and scientific standards. We posit that properly-controlled clinical trials are the most convincing means to determine the error rates of fMRI-based lie detection and confirm or disprove the relevance of the promising laboratory research on this topic.

This article explains the current state of the science and provides an analysis of the case law in which litigants have sought to introduce fMRI lie detection. Analyzing the myriad issues related to fMRI lie detection, the article identifies the key limitations of the current neuroimaging of deception science as expert evidence and explores the problems that arise from using scientific evidence before it is proven scientifically valid and reliable. We suggest that courts continue excluding fMRI lie detection evidence until this potentially useful form of forensic science meets the scientific standards currently required for adoption of a medical test or device.

Given a multitude of stakeholders and, the charged and controversial nature and the potential societal impact of this technology, goodwill and collaboration of several government agencies may be required to sponsor impartial and comprehensive clinical trials that will guide the development of forensic fMRI technology.

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 while 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.

The article seeks to explain for the interdisciplinary audience the pivotal difference between small-scale experimental research studies and properly controlled clinical trials that are dedicated to confirmation of the proofs of concept in the ecologically valid setting. 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 and speculative leap from existing data. Given the 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 are well known, a point mentioned in the National Academy of Sciences’ Report, *Strengthening Forensic Science in the United States: A Path Forward* (2009) (“NRC Report”). The NRC Report criticizes forensic science, questioning “whether—and to what extent—there is any science in any given ‘forensic science’ discipline.” It also finds the judiciary to be “utterly ineffective” in requiring forensic scientists to prove the validity of their methods and the accuracy of their conclusions (National Research Council, 2009). Despite the Report’s scathing critique of many forensic science disciplines, many courts continue to admit it without reservation, despite proof that such evidence has contributed to numerous wrongful convictions. Once admitted, scientific evidence tends to become rooted and difficult to eradicate later and we believe this report should influence the legal community to require 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 United States 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 shows 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, 2009; Stern, 2003).

The polygraph, which measures activity of the peripheral nervous system to gauge truthfulness, has been the primary technical method for lie detection during the last century. Beginning with the *Frye v. United States* (1923) decision, most United States courts have expressed disapproval of polygraph-based evidence. The United States Supreme Court has noted the lack of consensus on the reliability and admissibility of the polygraph (*Scheffer v. United States*, 1998), and courts remain largely hostile to its admission into evidence (Faigman, Kaye, Saks, & Sanders, 2010; Gallini, 2010). A meta-analysis commissioned by the 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% and 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 technique for government agencies, such

as the Federal Bureau of Investigation. Anecdotal evidence (Senate, 1994) and some retrospective studies 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 United States government workers, most of the 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 Functional Magnetic Resonance Imaging (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 fMRI was first reported in humans in 1992 (Kwong et al., 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 J.P. Rosenfeld, (Rosenfeld, Cantwell, Nasman, Wojdac, Ivanov, & Mazzeri, 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. The recent progress in the ability of fMRI to reliably measure and localize the activity of the central nervous system (CNS) 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.

a. BOLD fMRI and Principle of Cognitive Subtraction

Magnetic resonance imaging (MRI) is a medical imaging technique using high magnetic fields and non-ionizing electromagnetic radiation to produce high-resolution, three-dimensional (3D) tomographic images of the body (Lauterbur, 1973). Functional MRI (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 data (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). It follows that the selection of comparison conditions is essential to a meaningful experimental fMRI paradigm. Ideally, the comparison and target conditions would be identical except for a single factor 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 truth could be substituted for the two types of squares, but the same principle applies.

b. Experiment Design

fMRI deception experiment testing requires critical parameters, some of which are unique to fMRI and others that have been developed in basic psychological and polygraph research (Miller, 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 order of the stimuli presentation during an fMRI task (Donaldson & Buckner, 2001). In “event-related” paradigms, fMRI 3D images are acquired for discrete “events,” typically on the time course of one-half to four seconds. Event-related designs have an advantage in their ability to isolate activity in near-immediate response to stimuli and also allow for better stimulus variety and control types (Donaldson & Buckner, 2001). However, because of their low statistical power, they require random repetition of each class (i.e. lie or truth) of stimuli up to a dozen times during an experiment. Moreover, event-related designs require maximizing the magnetic field strength and the signal-to-noise ratio of the MRI scanner. Event-related or hybrid 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 have been 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 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 creates difficulty in adequately controlling 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 a piece 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). While not having the control problems of the CQT, the CIT’s reliance on the salience of

deception (rather than the deceptive response itself) limits its specificity. 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 parameter of importance to the experimental deception-generating models is whether responding deceptively is being *endorsed* by the experimenter (Miller, 1993). While in the real world, an individual's deception would generally be undesirable to its target (a feature known to the deceiver, by definition), in most deception experiments, subjects are given explicit instructions (i.e. endorsement) to lie to 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 to introduce intent 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 "co-conspirator" (Langleben et al. 2005).

c. Early Conclusions and Within-Subject Accuracy

Since 2000, academic researchers in several countries have used Blood Oxygenation Level Dependent (BOLD) functional Magnetic Resonance Imaging (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). While there remained inconsistencies across the multiple 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, 2010).

Simultaneously with the experimental progress, researchers recognized and explored the limitations and existing and potential pitfalls related to the possible translation of this technology to clinical use (Kozel, 2005). Wolpe, Foster and Langleben (2005) and Happel (2005) were the first to elaborate that a comprehensive understanding of the new technologies' error rates requires not only the recently reported within-subject accuracy, but also the positive and negative predictive power of the test, neither of which was known at the time (Hyman, 2010). The latter two parameters combine the inherent accuracy of a test and the expected prevalence of liars in the tested population and are a recognized milestone in the evaluation of clinical tests (Baldessarini, 1983). Their measurement requires large samples. Since 2005, this critical knowledge gap was underscored by several authoritative critics of the technology; however, no progress has been made in filling it.

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 (Marantz Henig, 2006; Talbot, 2007;).

Legal and ethical scholars also began to weigh in on the 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, Foster, & Langleben, 2005), constitutional implications (Fox 2009; Halliburton, 2009; Pardo, 2006) and privacy concerns (Greely, 2006; Happel, 2005; Thompson, 2005). Others suggested that while validation studies were necessary for translation of the fMRI lie detection into forensic practice, such studies were ethically and methodologically challenging. (Halber, 2007; Kanwisher, 2009). Halber argued that the accuracy rates of 80 to 90%, as reported in laboratory experiments, proved the method was inadequate for field applications (Halber, 2007).

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

France, however, 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 “[b]rain-imaging methods can be used only for medical or scientific research purposes or in the context of court expertise” (Oullier, 2012). 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).

fMRI Lie Detection in Court

Despite early, sustained criticism by both scientific and legal scholars, the 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, Inc., a company conducting commercial “credibility assessments” with fMRI.

Wilson v. Corestaff Services, L.P. 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.” 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:

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.

Wilson v. Corestaff Services L.P., (2010). 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*, 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 problems in the first two. The first session was negative for deception but deemed suboptimal. The second session was positive for deception but Semrau complained of fatigue during the scan. Apparently, the second session had excessively long test questions with double negatives, such as “Except for X, have you ever done Y?”. 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.

Magistrate Judge Tu M. Pham, appointed by the federal district court judge to hear the evidentiary motion, admitted testimony from opposing experts and reviewed affidavits submitted by experts. 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 (1) testability; (2) publication and peer review; (3) known error rate; (4) maintenance of standards and controls; and (5) 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, *i.e.* in the ‘real-world’ or ‘real-life’ setting;” a concern it voiced about both 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.” The court determined that the use of fMRI for deception in the real-world was not generally accepted by the scientific community and 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 (*United States v. Semrau*, 2011), providing an opportunity for the Court of Appeals for the Sixth Circuit 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 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.

There are four primary concerns these cases address and will likely be the focus of other courts' decisions as well: credibility, reliability, 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* 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 innocence.” We believe 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 disallow 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 up the jury's vital and exclusive function to make credibility determinations, and therefore does not ‘assist the trier of fact’ as required by FRE 702” (*United States v. Charley*, 1999). Not all courts, however, disfavor such testimony, and a minority of jurisdictions hold 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; Walker & Monahan, 1987). Much of this testimony helps the jury decide if 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 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, since 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 twentieth century, courts have been, at best,

skeptical of polygraph tests and, at worst and more usual, hostile to them.” (Faigman et al., at §40.1, 2010–2011); (Gallini, 2010). While such hostility may be due mostly polygraph’s limited reliability (Gallini, 2010; Scheffer, 1998; Stern, 2003), courts are concerned about invading the jury’s role (*United States v. Swayze*, 2004) and may be uncomfortable with technology that purports to know when people are lying.

The 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 is highly prejudicial where credibility is central to the verdict. (*United States v. Sherlin*, 1995). Nonetheless, discretion is granted to the 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 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 have a difficult time encouraging 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 and have not been tested either in real-life situations or in populations that deviate from what is 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. As we explained earlier, 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—the “task at hand”—and not globally (*Daubert*, 1993; 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, unrelated purposes” (*Daubert*, 1993). 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 precondition to admissibility.” (*Daubert*, 1993). Thus, under both a relevance and reliability analysis, fMRI evidence currently falls short of what is required for admissibility.

Third, courts that use the *Frye* general acceptance test (such as *Wilson* did) will also likely disfavor the evidence. While a few scientists on the advisory board of Cephus filed

affidavits in support of the science, most other neuroscientists involved with the fMRI lie detection research agree 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 multi-factor *Daubert* evidentiary reliability standard likewise uses the general acceptance factor in its analysis, and the lack of general acceptance among scientists in the field may well be critical to courts that follow *Daubert*.

Fourth, 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 evidence warrants (McCabe & Castel, 2008; Weisberg, Kiel, Goodstein, Rawson, & Gray, 2008), although there has been criticism of those studies (Schauer, 2010a). More recent data suggest 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. The authors conclude, that “the overwhelming consistent finding has been a lack of any impact of neuroimages on the decisions of our mock jurors” (Schweitzer et al., 2011). In the meta-analysis, the authors 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. While this study is compelling, there is more to be done in the area, a point well explained by the authors.

It is likely the Sixth Circuit will affirm the lower court’s decision in *Semrau*, since federal courts of appeals review lower court decisions about expert evidence under an abuse of discretion standard. (*General Electric Co. v. Joiner*, 1997). It is also unlikely the Court of Appeals will find that the trial court in *Semrau* abused its discretion in 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). The *Semrau* decision is well reasoned and well-grounded in both facts and science, and it is unlikely a court of appeals will overturn it. Even if *Semrau* is affirmed, however, the Court of Appeals may choose not to address the issue in depth, simply finding that the court below did not abuse its discretion. If that happens, then the Magistrate Judge’s opinion may not carry much weight with other courts, since 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 Court of Appeals may find that the defendant simply failed to meet the reliability standard in this case, but making no comment about 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 quite poor and likely worse than the fMRI evidence. 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. (Schauer, 2010b). If admitted, it should be as probabilistic rather than categorical evidence, much the way DNA evidence is admitted.

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 better experimentally grounded than the commonly admitted individualization evidence (fingerprints, handwriting, tool-marks, etc.) roundly criticized by the NRC Report (2009). Finally, other forms of neuroimaging, such as nuclear medicine (PET and 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).

Criminal defendants, however, may be able to introduce the evidence in certain types of proceedings. Due to constitutional rights, statutory enactments, and concerns over wrongful convictions, fMRI credibility assessment testimony might be admissible without meeting either the *Frye* or *Daubert* standards: in the penalty phases of capital cases, where defendants have a constitutional right to present mitigating evidence (*Smith v. Spivak*, 2010); or to support a claim of post-conviction 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 Supreme Court has consistently affirmed constitutional protections for defendants to introduce mitigating evidence in penalty hearings (*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. Kemp*, 1987). 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 than death" (*Penry v. Lynaugh*, 1989). A defendant may be able to make a compelling case that fMRI lie detection will meet this foregoing standard.

Although the Federal Rules of Evidence 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). Given that 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 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 in Illinois, *State v. Dugan*, the defense introduced expert testimony during the penalty phase that Dugan suffered from psychopathy, arguing

that it affected defendant's ability to control his impulse to kill (Hughes, 2010). The trial court allowed the expert to discuss the fMRI scans taken of Dugan's brain as additional proof of the defendant's psychopathy. The court also permitted expert testimony to help 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 did not permit him to use the actual fMRI images of Dugan's brain activity. Despite the admission of such expert testimony, Dugan received the death penalty. However, there was a signed verdict form discovered after the sentencing indicating that the jury actually 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 post-trial 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. Upon 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 upon 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 upon 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% to 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 out 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 out of 15 were correctly identified, yielding 100% sensitivity but only 33% specificity.

Another study using a within-subject design and a sophisticated non-parametric analysis (Davatzikos, 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 countermeasures. 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.

Importantly, the group differences between lie and truth consistently involved the lateral and inferior prefrontal and posterior parietal cortices and appear unaffected by gender, handedness and language. While 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 (1) stakes are higher; (2) 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 (3) the effects of motor and cognitive countermeasures are evaluated in a deliberate fashion.

Finally, while 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, Rosenfeld, Meixner, et al., 2011; Kaylor-Hughes, et al., 2010; Kozel, et al., 2009; Mildner, Zysset, Trampel, Driesel, & Moller, 2005; Nunez, 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, 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, BOLD fMRI describes changes in regional brain activity over time rather than providing 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.* 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 pre-set 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 vs. distributed functions in the brain that dates back to the nineteenth century and may continue well after the determination of the utility and scope of the potential use of fMRI for lie

detection is complete. As an analogy, we use antidepressant drugs extensively, without knowing their exact mechanisms of action.

Though the basic cognitive neuroscience study of deception is clinically important in the long run, the critical question of error rates and other translational questions described earlier can and should be answered independently of the basic research on the mechanisms, since 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 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 (Shen and Jones, 2011). 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).

While 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 to say, whether the test actually measures what it purports to. 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 indeed 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 notes that the court has the discretion “to choose among *reasonable* means of excluding expertise that is *fausse* and science that is junky.”

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, the Supreme Court comments 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—*i.e.*, ‘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 reliability.” In footnote 9 following the quote, the court explains 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* will be based upon *scientific validity*.

Thus the court defines “legal reliability” in terms of “scientific validity.” While 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 71% to greater than 90% accuracy. Is that enough for “legal reliability”? We do not think so. Without knowing the positive and negative predictive power of the test, there is no accurate way to respond to Daubert’s “known error rate” inquiry. This science is currently in 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*, 1997). 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 evidence (Cutler and 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, including functional brain imaging” (Raichle, 2009).

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 examined the trial transcripts of 137 exonerated defendants and concluded that approximately 60% of those trials included flawed science (Garrett and Neufeld, 2009). The NRC Report concludes 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” (NRC, 2009). These forms of forensic evidence include fingerprints, toolmarks, handwriting, bitemarks and hair comparison; the often critical evidence in criminal trials.

The Report 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 or the problems of both bias and variability among those performing the analysis and no rigorous protocols to guide the subjective interpretations. The 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.” (NRC, 2009). Despite the scathing 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. (*United States v. Love*, 2011) (fingerprint comparison); (*Pettus v. United States*, 2012) (handwriting).

To date, 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-review articles on the subject establish. Yet we do not believe that even if fMRI lie detection is better than much of forensic science 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

Though fMRI lie detectors are not ready for legal application, we 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, 2009). Objective means for detecting deception have a high potential 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 non-adherence in chronic medical illness. Thus, the topic is well worth pursuing with both translational forensic 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 polygraph, there has been no major movement to legislatively ban its use in all circumstances, except for non-Government pre-employment testing (OTA, 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 707, which bans polygraph evidence, despite the Supreme Court upholding the constitutionality of MRE 707 (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 demographic and the prevalence of the behavior in question.

Clinical validation of a test is an expensive enterprise usually 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 what is likely to require the large overall number of participants, though the number required for each study could be relatively small (50–100). 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's example, the predictive value of a positive test (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 1000 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 could cost between 5 and 15 million dollars, a price tag below an average pharmaceutical company study. 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 vaccines for drug addiction, the United States 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 non-adherence, so far NIH has not recognized deception as a health issue. United States defense and intelligence agencies have funded research in this area, but its results have been slow to appear in scientific literature (Dedman, 2009; Moreno, 2006; Stern, 2003) and may be subject to non-scientific bias similar to those that afflicted the 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 \$10M Law and Neuroscience Project 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 project outside of the purview of the National Science Foundation.

Conclusion

In 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 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 a multitude of stakeholders, the charged and 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.

Acknowledgments

The authors wish to thank Kenneth Foster of the University of Pennsylvania for helpful comments on an early draft; the anonymous reviewers for their helpful comments throughout the peer-review process; Christy Gamble and Gabrielle Steier of the Duquesne University School of Law for research assistance; and Bridget Praskovich for editorial and formatting assistance.

References

- Abe N, Fujii T, Hirayama K, Takeda A, Hosokai Y, Ishioka T, Mori E. Do parkinsonian patients have trouble telling lies? The neurobiological basis of deceptive behaviour. *Brain*. 2009; 132(Pt 5):1386–1395.10.1093/brain/awp052 [PubMed: 19339257]
- Aguirre, GK.; D’Esposito, M. Experimental design for brain fMRI. In: Moonen, CTW.; Bandettini, PA., editors. *Functional MRI*. New York: Springer; 1999. p. 369-380.
- Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ*. 1995; 311(7003): 485.10.1136/bmj.311.7003.485 [PubMed: 7647644]
- Baldessarini RJ, Finklestein S, Arana GW. The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry*. 1983; 40:569–573. [PubMed: 6838334]
- Barnum, A.; St Clair, S. Dugan juror who switched at last minute stands by death sentence. *Chicago Tribune*. 2009. http://articles.chicagotribune.com/2009-12-03/news/0912021217_1_death-sentence-two-jurors-death-penalty
- Barnum, A. Dugan asks to drop appeal in Nicarico case. *Chicago Tribune*. 2009. Retrieved from http://articles.chicagotribune.com/2011-03-11/news/ct-met-dugan-appeal-0312-20110311_1_brian-dugan-jeanine-nicarico-nicarico-family
- Belliveau JW, Kwong KK, Kennedy DN, Baker JR, Stern CE, Benson R, et al. Magnetic resonance imaging mapping of brain function. *Human visual cortex*. *Investigative Radiology*. 1992; 27(Suppl 2):S59–65. [PubMed: 1468876]
- Ben-Shakhar G. Clinical judgment and decision-making in CQT-polygraphy. A comparison with other pseudoscientific applications in psychology. *Integrative Physiological & Behavioral Science*. 1991; 26:232–240.10.1007/BF02912515 [PubMed: 1954163]
- Ben-Shakhar, G. A critical review of the Controlled Question Test (CQT). In: Kleiner, M., editor. *Handbook of polygraph testing*. London and San Diego: Academic Press; 2001. p. 103-127.
- Ben-Shakhar G, Bar-Hillel M, Kremnitzer M. Trial by polygraph: Reconsidering the use of the guilty knowledge technique in court. *Law & Human Behavior*. 2002; 26:527–541.10.1023/A: 1020204005730 [PubMed: 12412496]
- Berger H. Über das Elektroencephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*. 1929; 87:527–570.
- Bizzi, E.; Hyman, S.; Raichle, M.; Kanwisher, N.; Phelps, E.; Morse, S., et al. Using imaging to identify deceit: Scientific and ethical questions. Washington, DC: 2009.

- Canli T, Brandon S, Casebeer W, Crowley PJ, Du Rousseau D, Greely HT, et al. Neuroethics and national security. *American Journal of Bioethics*. 2007; 7(5):3–13.10.1080/15265160701290249 [PubMed: 17497494]
- Crewson, PE. Comparative analysis of polygraph with other screening and diagnostic tools. Ashburn, VA: Research Support Service; 2001. (No. DODPI00P0030 (DODPI00P0030) XDDODPI/SC (XDDODPISC))
- Cutler, BL.; Wells, GL. Psychological science in the courtroom: Consensus and controversy. In: Skeem, JL.; Douglas, KS., editors. *Expert testimony regarding eyewitness identification*. New York: Guilford Press; 2009. p. 100-123.
- 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 DG, Acharyya M, Loughead JW, et al. Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *Neuroimage*. 2005; 28:663–668.10.1016/j.neuroimage.2005.08.009 [PubMed: 16169252]
- Dedman, B. New anti-terror weapon: Hand-held lie detector. 2008. Retrieved [3.11.2012], from <http://www.msnbc.msn.com/id/23926278/>
- Donaldson, DI.; Buckner, RL. *Functional MRI: An introduction to methods*. New York, NY: Oxford University Press; 2001. Effective paradigm design; p. 176-195.
- Ekman P, O'Sullivan M. Who can catch a liar? *American Psychologist*. 1991; 46:913–920.10.1037/0003-066X.46.9.913 [PubMed: 1958011]
- Faigman DL, Kaye DH, Saks MJ, Sanders J 2. *Mod Sci Evi*. §17(10):2010–2011.
- Faigman, DL.; Kaye, DH.; Saks, MJ.; Sanders, J. Modern scientific evidence. 2010. §40.1 Polygraph tests, legal issues
- Fisher G. The jury's rise as lie detector. *Yale Law Journal*. 1997; 107:575–713.
- Frye v. United States, 293 F. 1013 (D.C. Cir. 1923)
- Fullam RS, McKie S, Dolan MC. Psychopathic traits and deception: functional magnetic resonance imaging study. *British Journal of Psychiatry*. 2009; 194:229–235.10.1192/bjp.bp.108.053199 [PubMed: 19252152]
- Gallini BR. Police science in the interrogation room: Seventy years of pseudo-psychological interrogation methods to obtain inadmissible confessions. *Hastings Law Journal*. 2010; 61(26): 529–580. 534.
- Ganis G, Kosslyn SM, Stose S, Thompson WL, Yurgelun-Todd DA. Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex*. 2003; 13:830–836.10.1093/cercor/13.8.830 [PubMed: 12853369]
- Ganis G, Rosenfeld JP, Meixner J, Kievit RA, Schendan HE. Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *Neuroimage*. 2011; 55:312–319.10.1016/j.neuroimage.2010.11.025 [PubMed: 21111834]
- Garrett BL, Neufeld PJ. Invalid forensic science testimony and wrongful conviction. *Virginia Law Review*. 2009; 95:1–76.
- Gazzaniga MS. The law and neuroscience. *Neuron*. 2008; 60:412–415.10.1016/j.neuron.2008.10.022 [PubMed: 18995814]
- Gen. Elec. Co. v. Joiner, 522 U.S. 136, 118 S. Ct. 512, 139 L. Ed. 2d 508 (1997)
- Gjedde, A. Brain energy metabolism and the physiological basis of the haemodynamic response. In: Jezzard, P.; Matthews, PM.; Smith, S., editors. *Functional MRI: An introduction to methods*. New York, NY: Oxford University Press; 2001. p. 37-66.
- Greely H. To tell the truth. Brain scans should not be used for lie detection unless their reliability is proven. *Scientific American*. 2010; 303:18.10.1038/scientificamerican1210-18 [PubMed: 21141350]
- Greely HT, Illes J. Neuroscience-based lie detection: The urgent need for regulation. *American Journal of Law & Medicine*. 2007; 33:377–431. [PubMed: 17910165]
- Greely, HT. The social effects of advances in neuroscience: Legal problems, legal perspective. In: Illes, J., editor. *Neuroethics: Defining the issues in theory, practice and policy*. New York: Oxford University Press; 2006. p. 245-263.

- Greene JD, Paxton JM. Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:12506–12511.10.1073/pnas.0900152106 [PubMed: 19622733]
- Griffin L. Correcting Injustice: How the United Kingdom and the United States review claims of innocence. *Toledo Law Review*. 2009; 41:107–169.
- Grubin D. The polygraph and forensic psychiatry. *Journal of the American Academy of Psychiatry & the Law*. 2010; 38:446–451. [PubMed: 21156902]
- Hakun JG, Ruparel K, Seelig D, Busch E, Loughhead JW, Gur RC, et al. Towards clinical trials of lie detection with fMRI. *Society for Neuroscience*. 2009; 4:518–527.10.1080/17470910802188370
- Halber, D. Scientists swear a good lie detector is hard to find. 2007 Feb 14. Retrieved from MIT, MIT Tech Talk website: <http://web.mit.edu/newsoffice/2007/techtalk51-17.pdf>
- Happel MD. Neuroscience and the detection of deception. *Review of Policy Research*. 2005; 22:667–685.
- Harrington v. State, 659 N.W.2d 509 (Iowa 2003).
- Hickman, W. Senate approves \$100 million for South Carolina defense projects. 2005. Retrieved November 1, 2008, from <http://lgraham.senate.gov/public/>
- Hughes V. Science in court: Head case. *Nature*. 2010; 464(7287):340–342.10.1038/464340a [PubMed: 20237536]
- Hyman SE. 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*. 2010; 10:49–50.10.1080/15265161.2010.527263 [PubMed: 20945267]
- Imwinkelried EJ. 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*. 2011; 62:959–987.
- Kanwisher, N. Using imaging to identify deceit: Scientific and ethical questions. Cambridge, MA: American Academy of Arts & Sciences; 2009. The use of fMRI in lie detection: What has been shown and what has not; p. 7-13.
- Kaye, DH.; Bernstein, DE.; Mnookin, JL. The new Wigmore: A treatise on evidence: Expert evidence. 2. 2012. §2.4 Expert testimony on credibility: General rule; p. 1-4. Retrieved from Westlaw database
- Kaylor-Hughes CJ, Lankappa ST, Fung R, Hope-Urwin AE, Wilkinson ID, Spence SA. The functional anatomical distinction between truth telling and deception is preserved among people with schizophrenia. *Criminal Behaviour & Mental Health*. 2010; 21:8–20.10.1002/cbm.785 [PubMed: 20661881]
- Kozel FA, Johnson KA, Grenesko EL, Laken SJ, Kose S, Lu X, et al. Functional MRI detection of deception after committing a mock sabotage crime. *Journal of Forensic Sciences*. 2009; 54:220–231.10.1111/j.1556-4029.2008.00927.x [PubMed: 19067772]
- Kozel FA, Johnson KA, Mu Q, Grenesko EL, Laken SJ, George MS. Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*. 2005; 58:605–613.10.1016/j.biopsych.2005.07.040 [PubMed: 16185668]
- Kozel FA, Trivedi MH. Developing a neuropsychiatric functional brain imaging test. *Neurocase*. 2008; 15(1):1–5.10.1080/13554790701881731 [PubMed: 19065284]
- Kumho Tire Co., Ltd. v. Carmichael, 526 U.S. 137, 119 S. Ct. 1167, 143 L. Ed. 2d 238 (1999).
- Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *PNAS*. 1992; 89:5675–5679.10.1073/pnas.89.12.5675 [PubMed: 1608978]
- Langleben DD. Detection of deception with fMRI: Are we there yet? *Legal & Criminological Psychology*. 2008; 13:1–9.10.1348/135532507X251641
- Langleben DD, Loughhead JW, Bilker WB, Ruparel K, Childress AR, Busch SI, et al. Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*. 2005; 26:262–272.10.1002/hbm.20191 [PubMed: 16161128]

- Langleben DD, Schroeder L, Maldjian JA, Gur RC, McDonald S, Ragland JD, et al. Brain activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage*. 2002; 15:727–732.10.1006/nimg.2001.1003 [PubMed: 11848716]
- Langleben, DD.; Willard, D.; Moriarty, JC. Detecting Deception. In: Simpson, JR., editor. *Neuroimaging in forensic psychiatry: From the clinic to the courtroom*. Chichester: Wiley-Blackwell; 2012.
- Lauterbur PC. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*. 1973; 242(5394):190–191.10.1038/242190a0
- Lee TM, Liu HL, Tan LH, Chan CC, Mahankali S, Feng CM, et al. Lie detection by functional magnetic resonance imaging. *Human Brain Mapping*. 2002; 15:157–164.10.1002/hbm.10020 [PubMed: 11835606]
- Lykken DT. Why (some) Americans believe in the lie detector while others believe in the guilty knowledge test. *Integrative Physiological & Behavioral Science*. 1991; 26:214–222.10.1007/BF02912513 [PubMed: 1954161]
- Lynch WJ, Roth ME, Carroll ME. Biological basis of sex differences in drug abuse: Preclinical and clinical studies. *Psychopharmacology (Berl)*. 2002; 164:121–137.10.1007/s00213-002-1183-2 [PubMed: 12404074]
- Marantz Henig, R. *New York Times*. 2006 Feb 6. Looking for the lie: The new science of lying. Brain mapping, E.R.P. waves and microexpression readings could make deception harder. But would that make us better?; p. 47
- Melton, GB.; Petrila, J.; Poythress, NG.; Slobogin, C. *Psychological evaluations for the courts: A handbook for mental health professional and lawyers*. 3. New York: Guilford Press; 2007.
- Mildner T, Zysset S, Trampel R, Driesel W, Moller HE. Towards quantification of blood-flow changes during cognitive task activation using perfusion-based fMRI. *Neuroimage*. 2005; 27:919–926.10.1016/j.neuroimage.2005.04.040 [PubMed: 15978840]
- Miller, GRS.; JB. *Deceptive Communication*. Vol. 14. Newbury Park, CA: Sage; 1993. Investigating deceptive communication; p. 32-39.
- Monahan J, Walker L. Social frameworks: A new use of social science in law. *Virginia Law Review*. 1987; 73:559–598.
- Monahan J, Walker L, Mitchell G. Contextual evidence of gender discrimination: The ascendance of “social frameworks. *Virginia Law Review*. 2008; 94:1715–1748.
- Moriarty JC. Flickering admissibility: Neuroimaging evidence in the U.S. courts. *Behavioral Sciences & the Law*. 2008; 26:29–49.10.1002/bsl.795 [PubMed: 18327830]
- Moriarty JC. Visions of deception: Neuroimages and the search for truth. *Akron Law Review*. 2009; 42:739–762.
- Moriarty JC. Will history be servitude? The NAS Report on forensic science and the role of the judiciary. *Utah Law Review*. 2010; 2010:299–326.
- Myers JEB. Expert testimony in child sexual abuse litigation: Consensus and confusion. *University of California Davis Journal of Juvenile Law & Policy*. 2010; 14:1–57.
- Nakayama, M. Practical of the concealed information test for criminal investigations in Japan. In: Kleiner, M., editor. *Handbook of polygraph testing*. London and San Diego: Academic Press; 2001. p. 49-86.
- Nance DA, Morris SB. An empirical assessment of presentation formats with a relatively large and quantifiable random match probability. *Jurimetrics Journal*. 2002; 42:403–445.
- Nance DA, Morris SB. 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*. 2005; 34:395–443.
- National Research Council . *Strengthening forensic science in the United States: A path forward*. Washington, D.C: National Academy of the United States; 2009.
- Nunez JM, Casey BJ, Egner T, Hare T, Hirsch J. Intentional false responding shares neural substrates with response conflict and cognitive control. *Neuroimage*. 2005; 25:267–277.10.1016/j.neuroimage.2004.10.041 [PubMed: 15734361]
- OTA. *The use of integrity tests for pre-employment screening*. Washington, DC: U.S. Office of Technology Assessment; 1990. (No. Report No. OTA-SET-442)

- Oullier O. Clear up this fuzzy thinking on brain scans. *Nature*. 2012 Mar 1;483:7.10.1038/483007a [PubMed: 22382941]
- Owen, AM.; Epstein, R.; Johnsrude, IS. fMRI: Applications in cognitive neuroscience. In: Jezzard, P.; Matthews, PM.; Smith, S., editors. *Functional MRI: An introduction to methods*. New York, NY: Oxford University Press; 2001. p. 311-327.
- Pardo MS. Neuroscience evidence, legal culture, and criminal procedure. *American Journal of Criminal Law*. 2006; 33:301, 319.
- Pettus v. United States, ___A.3d.___, 2012 WL 399997 (D.C. 2012).
- Phelps, EA. Using imaging to identify deceit: Scientific and Ethical Questions. *American Academy of Arts & Sciences*; 2009. Lying outside the laboratory: The impact of imagery and emotion on the neural circuitry of lie detection; p. 14-22.
- Poulin AB. Credibility: A fair subject for expert testimony? *Florida Law Review*. 2007; 59:991–1069.
- Raichle, ME. Using imaging to identify deceit: Scientific and ethical questions. *American Academy of Arts & Sciences*; 2009. An introduction to functional brain imaging in the context of lie detection; p. 3-6.
- Rakoff, JS. Using imaging to identify deceit: Scientific and ethical questions. *CITY: American Academy of Arts & Sciences*; 2009. Lie detection in the courts: The vain search for the magic bullet; p. 40-45.
- Raskin, DC.; Honts, CR. The comparison question test. In: Kleiner, M., editor. *Handbook of polygraph testing*. London and San Diego: Academic Press; 2001. p. 14-15.
- Risinger DM. Defining the “task at hand”: Non-science forensic science after *Kumho Tire Co. v. Carmichael*. *Washington & Lee Law Review*. 2000; 57:767–800.
- Rosenfeld JP, Cantwell B, Nasman VT, Wojdac V, Ivanov S, Mazzeri L. A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience*. 1988; 42:157–161.10.3109/00207458808985770 [PubMed: 3209369]
- Rosenfeld JP. “Brain fingerprinting”: A critical analysis. *The Scientific Review of Mental Health Practice*. 2005; 4(1):20–37.
- Rushing, SE.; Pryma, DA.; Langleben, DD. PET and SPECT. In: Simpson, JR., editor. *Neuroimaging in Forensic Psychiatry: From the clinic to the courtroom*. Chichester: Wiley-Blackwell; 2012.
- Schauer F. Can bad science be good evidence? *Cornell Law Review*. 2010a; 95:1191–1220. [PubMed: 20939147]
- Schauer F. 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*. 2010b; 14:101–103.10.1016/j.tics.2009.12.004 [PubMed: 20060772]
- Schweitzer NJ, Saks MJ, Murphy ER, Roskies AL, Sinnott-Armstrong W, Gaudet LM. Neuroimages as evidence in a mens rea defense: No impact. *Psychology, Public Policy, and Law*. 2011; 17:357–393.10.1037/a0023581
- Seaman JA. Black boxes. *Emory Law Journal*. 2008; 58:427–488.
- U. S. Senate. An assessment of the Aldrich H. Ames espionage case and its implications for U.S. intelligence. Washington, DC: U. S. Congress; 1994.
- Simpson JR. Functional MRI lie detection: Too good to be true? *Journal of the American Academy of Psychiatry & the Law*. 2008; 36:491–498. [PubMed: 19092066]
- Smith v. Spisak, 130 S. Ct. 676, 175 L. Ed. 2d 595 (2010)
- Spence SA. Playing devil’s advocate: The case against fMRI lie detection. [Editorial]. *Legal and Criminological Psychology*. 2008; 13:11–25.10.1348/135532507X251597
- Spence SA, Farrow TF, Herford AE, Wilkinson ID, Zheng Y, Woodruff PW. Behavioural and functional anatomical correlates of deception in humans. *Neuroreport*. 2001; 12:2849–2853.10.1097/00001756-200109170-00019 [PubMed: 11588589]
- Spence SA, Hunter MD, Farrow TF, Green RD, Leung DH, Hughes CJ, et al. A cognitive neurobiological account of deception: Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*. 2004; 359(1451): 1755–1762.10.1098/rstb.2004.1555

- Spence SA, Kaylor-Hughes CJ. Looking for truth and finding lies: The prospects for a nascent neuroimaging of deception. *Neurocase*. 2008; 14:68–81.10.1080/13554790801992776 [PubMed: 18569733]
- Spence SA, Kaylor-Hughes CJ, Brook ML, Lankappa ST, Wilkinson ID. ‘Munchausen’s syndrome by proxy’ or a ‘miscarriage of justice’? An initial application of functional neuroimaging to the question of guilt versus innocence. *Eur Psychiatry*. 2008; 23(4):309–314.10.1016/j.eurpsy.2007.09.001 [PubMed: 18029153]
- Spence SA, Hope-Urwin A, Lankappa ST, Woodhead J, Burgess JCL, Mackay AV. If brain scans really detected deception, who would volunteer to be scanned? *Journal of Forensic Science*. 2010; 55(5):1352–1355.10.1111/j.1556-4029.2010.01452.x
- Stern, PC., editor. *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; 2003.
- Talbot, M. Duped: Can brain scans uncover lies?. *The New Yorker*. 2007. Retrieved [date], from http://www.newyorker.com/reporting/2007/07/02/070702fa_fact_talbot
- The Law and Neuroscience Project. 2010. Retrieved June 29, 2011, from <http://www.lawneuro.org/>
- Thompson SK. The legality of the use of psychiatric neuroimaging in intelligence interrogation. *Cornell Law Review*. 2005; 90:1601–1637. [PubMed: 16220619]
- Tovino SA. Imaging body structure and mapping brain function: A historical approach. *American Journal of Law and Medicine*. 2007; 33:193–228. [PubMed: 17910157]
- United States v. Charley, 189 F. 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. Smith, 405 Fed. App’x 795 (4th Cir. 2010).
- United States v. Swayze, 378 F.3d 834, 837 (8th Cir. 2004).
- Vrij, A. *Detecting lies and deceit: Pitfalls and opportunities*. 2. Chichester, England: Wiley; 2008.
- Wagner A. Can neuroscience identify lies? A Judge’s Guide to Neuroscience: A Concise Introduction. 2010:13–25.
- Weisberg DS, Keil FC, Goodstein J, Rawson E, Gray JR. The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*. 2008; 20(3):470–477.10.1162/jocn.2008.20040 [PubMed: 18004955]
- Wilson v. Corestaff Services, L.P., 28 Misc. 3d 425, 900 N.Y.S.2d 639 (N.Y. Sup. Ct. 2010).
- Wolpe PR, Foster KR, Langleben DD. Emerging neurotechnologies for lie-detection: Promises and perils. *American Journal of Bioethics*. 2005; 5:39–49.10.1080/15265160590923367 [PubMed: 16036700]
- Yang Y, Raine A, Narr KL, Lencz T, LaCasse L, Colletti P, et al. Localisation of increased prefrontal white matter in pathological liars. *British Journal of Psychiatry*. 2007; 190:174–175.10.1192/bjp.bp.106.025056 [PubMed: 17267937]