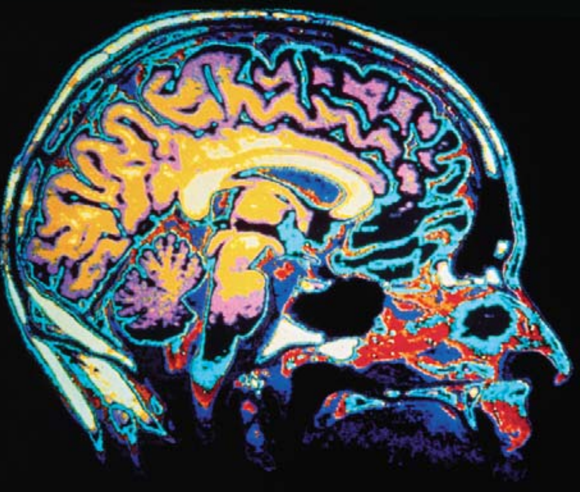
Using Imaging to Identify Deceit



Scientific and Ethical Questions

Emilio Bizzi, Steven E. Hyman, Marcus E. Raichle,
Nancy Kanwisher, Elizabeth A. Phelps, Stephen J. Morse,
Walter Sinnott-Armstrong, Jed S. Rakoff, and
Henry T. Greely

CHERISHING KNOWLEDGE · FOUNDED 1780 · SHAPING THE FUTURE
AMERICAN ACADEMY OF ARTS & SCIENCES

Using Imaging to Identify Deceit: Scientific and Ethical Questions

Emilio Bizzi, Steven E. Hyman, Marcus E. Raichle,
Nancy Kanwisher, Elizabeth A. Phelps,
Stephen J. Morse, Walter Sinnott-Armstrong,
Jed S. Rakoff, and Henry T. Greely

ISBN#: 0-87724-077-9

The views expressed in this volume are those held by each contributor and are not necessarily those of the Officers and Fellows of the American Academy of Arts and Sciences.

Please direct inquiries to:
American Academy of Arts and Sciences
136 Irving Street
Cambridge, MA 02138-1996
Telephone: (617) 576-5000
Fax: (617) 576-5050
Email: aaas@amacad.org
Web: www.amacad.org

Contents

Imaging Deception

EMILIO BIZZI AND STEVEN E. HYMAN

Can the relatively new technique of functional magnetic resonance imaging (fMRI) detect deceit? A symposium sponsored by the American Academy of Arts and Sciences, the McGovern Institute at the Massachusetts Institute of Technology (MIT), and Harvard University took on this question by examining the scientific support for using fMRI as well as the legal and ethical questions raised when machine-based means are employed to identify deceit.

Marcus Raichle, a professor at Washington University in St. Louis, opens the discussion with a clear description of fMRI, its physiological basis, the methodology underlying the extraction of images, and, most important, the use of image averaging to establish correlations between the "images" and aspects of behavior. While averaging techniques are highly effective in the characterization of functional properties of different brain areas, images obtained from a single individual are "noisy," a fact that clearly touches upon the reliability of the extracted data and a fortiori makes detecting deceit a questionable affair.

Nancy Kanwisher, a professor at MIT, discusses papers that present supposedly direct evidence of the efficacy of detecting deceit with fMRI, but dismisses their conclusions. Kanwisher notes that there is an insurmountable problem with the experimental design of the studies she analyzes. She points out that by necessity the tested populations in the studies consisted of volunteers, usually cooperative students who were asked to lie. For Kanwisher this experimental paradigm bears no relationship to the real-world situation of somebody brought to court and accused of a serious crime.

Kanwisher's conclusions are shared by Elizabeth Phelps, a professor at New York University. Phelps points out that two cortical regions—the parahippocampal cortex and the fusiform gyrus—display different activity in relation to familiarity. The parahippocampal cortex shows more activity for less familiar faces, whereas the fusiform gyrus is more active for familiar faces. However, these neat distinctions can unravel when imagined memories are generated by subjects involved in emotionally charged situations. Phelps points out that the brain regions important to memory do not differentiate between imagined memories and those based on events in the real world. In addition, the perceptual details of memories are affected by emotional states.

Phelps's compelling description of how imagination, emotions, and misperceptions all play a role in shaping memories can be briefly expressed as "brains do not lie: people do." This point is echoed by Stephen Morse, who begins his presentation by stating "Brains do not commit crimes. Acting people do." Morse, a professor at the University of Pennsylvania, takes a skeptical view of the potential contributions of neuroscience in the courtroom. He believes that behavioral evidence is usually more useful and informative than information based on brain science, and that when neuroimaging data and behavioral evidence conflict, the behavioral evidence trumps imaging. Morse worries that admitting imaging in the courtroom might sway "naive" judges and jurors to think that the brain plays a "causal" role in a crime. He repeatedly warns that if causation excuses behavior then no one can ever be considered responsible.

Walter Sinnott-Armstrong, a professor at Dartmouth College, is also unenthusiastic about the use of fMRI to detect deceit. His concern is that the error rates in fMRI are significant and that determining error rates is not a simple task. For this reason he believes that evidence from neural lie detection efforts should not be allowed in court.

Jed Rakoff, a U.S. district judge, shares Sinnott-Armstrong's concern about error rates and finds that fMRI-based evidence may be excluded from trials under the Federal Rules of Evidence. Rakoff argues that the golden path to discovering truth is the traditional one of exposing witnesses to cross-examination. He doubts that meaningful correlations between lying and brain images can be reliably established. In addition, he notes that the law recognizes many kinds of lies—for example, lies of omission, "white lies," and half-truths—and asks whether brain imaging can come close to distinguishing among these complex behavioral responses. Clearly not, he concludes, but traditional cross-examination might do the job.

Henry Greely, a professor at Stanford Law School, discusses the constitutional and ethical issues raised by fMRI lie detection. He cites as concerns the problems related to the scientific weakness of some fMRI studies, the disagreement among the investigators about which brain regions are associated with deception, the limitations of pooled studies, and the artificiality of experimental design.

The authors of these seven essays express a dim view of lie detection with fMRI. They also consider the widely used polygraph and conclude that both it and fMRI are unreliable.

Often in science when a new technique such as fMRI appears, the scientists who promote its use argue that, yes, problems exist but more research will, in the end, give us the magic bullet. Perhaps. In the case of lie detection through fMRI, however, two sets of problems seem insurmountable: 1) problems of research design, which Kanwisher argues no improvement in imaging technology is likely to address; and 2) problems of disentangling emotions, memory, and perception, which, Phelps notes, are processed in the same region of the brain and thus are commingled.

CHAPTER 1

An Introduction to Functional Brain Imaging in the Context of Lie Detection

MARCUS E. RAICHLE

Human brain imaging, as the term is understood today, began with the introduction of X-ray computed tomography (i.e., CT as it is known today) in 1972. By passing narrowly focused X-ray beams through the body at many different angles and detecting the degree to which their energy had been attenuated, Godfrey Hounsfield was able to reconstruct a map of the density of the tissue in three dimensions. For their day, the resultant images of the brain were truly remarkable. Hounsfield's work was a landmark event that radically changed the way medicine was practiced in the world; it spawned the idea that three-dimensional images of organs of the body could be obtained using the power of computers and various detection strategies to measure the state of the underlying tissues of the body.

In the laboratory in which I was working at Washington University in St. Louis, the notion of positron emission tomography (PET) emerged shortly after the introduction of X-ray computed tomography. Instead of passing an X-ray beam through the tissue and looking at its attenuation as was done with X-ray computed tomography, PET was based on the idea that biologically important compounds like glucose and oxygen labeled with cyclotron-produced isotopes (e.g., ^{15}O, ^{11}C, and ^{18}F) emitting positrons (hence the name *positron emission tomography*) could be detected in three dimensions by ringing the body with special radiation detectors. The maps arising from this strategy provided us with the first quantitative maps of brain blood flow and metabolism, as well as many other interesting measurements of function. With PET, modern human brain imaging began measuring function.

In 1979, magnetic resonance imaging (MRI) was introduced. While embracing the concept of three-dimensional imaging, this technique was based on the magnetic properties of atoms (in the case of human imaging, the primary atom of interest has been the hydrogen atom or proton). Studies of these properties had been pursued for several decades in chemistry laboratories using a technique called nuclear magnetic resonance. When this technique was applied to the human body and images began to emerge, the name was changed to "magnetic resonance imaging" to assuage concerns about radioactivity that might mistakenly arise because of the use of the term *nuclear.* Functional MRI (fMRI) has become the dominant mode of imaging function in the human brain.

At the heart of functional brain imaging is a relationship between blood flow to the brain and the brain's ongoing demand for energy. The brain's voracious appetite for energy derives almost exclusively from glucose, which in the brain is broken down to carbon dioxide and water. The brain is dependent on a continuing supply of both oxygen and glucose delivered in flowing blood regardless of moment-to-moment changes in an individual's activities.

For over one hundred years scientists have known that when the brain changes its activity as an individual engages in various tasks the blood flow increases to the areas of the brain involved in those tasks. What came as a great surprise was that this increase in blood flow is accompanied by an increase in glucose use but not oxygen consumption. As a result, areas of the brain transiently increasing their activity during a task contain blood with increased oxygen content (i.e., the supply of oxygen becomes greater than the demand for oxygen). This observation, which has received much scrutiny from researchers, paved the way for the introduction of MRI as a functional brain tool.

By going back to the early research of Michael Faraday in England and, later, Linus Pauling in the United States, researchers realized that hemoglobin, the molecules in human red blood cells that carry oxygen from the lungs to the tissue, had interesting magnetic properties. When hemoglobin is carrying a full load of oxygen, it can pass through a magnetic field without causing any disturbance. However, when hemoglobin loses oxygen to the tissue, it disrupts any magnetic field through which it passes. MRI is based on the use of powerful magnetic fields, thousands of times greater than the earth's magnetic fields. Under normal circumstances when blood passes through an organ like the brain and loses oxygen to the tissue, the areas of veins that are draining oxygen-poor blood show up as little dark lines in MRI images, reflecting the loss of the MRI signal in those areas. Now suppose that a sudden increase in blood flow locally in the brain is not accompanied by an increase in oxygen consumption. The oxygen content of these very small draining veins increases. The magnetic field in the area is restored, resulting in a local increase in the imaging signal. This phenomenon was first demonstrated with MRI by Seiji Ogawa at Bell Laboratories in New Jersey. He called the phenomenon the "blood oxygen level dependent" (BOLD) contrast of MRI and advocated its use in monitoring brain function. As a result researchers now have fMRI using BOLD contrast, a technique that is employed thousands of times daily in laboratories throughout the world.

A standard maneuver in functional brain imaging over the last twenty-five years has been to isolate changes in the brain associated with particular tasks by subtracting images taken in a control state from the images taken during the performance of the task in which the researcher is interested. The control state is often carefully chosen so as to contain most of the elements of the task of interest save that which is of particular interest to the researcher. For example, to "isolate" areas of the brain concerned with reading words aloud, one might select as the control task passively viewing words. Having eliminated areas of the brain concerned with visual word perception, the resulting "difference image" would contain only those areas concerned with reading aloud.

Another critical element in the strategy of functional brain imaging is the use of image averaging. A single difference image obtained from one individual appears "noisy," nothing like the images usually seen in scientific articles or the popular press. Image averaging is routinely applied to imaging data and usually involves averaging data from a group of individuals. While this technique is enormously powerful in detecting common features of brain function across people, in the process it completely obscures important individual differences. Where individual differences are not a concern, this is not a problem. However, in the context of lie detection researchers and others are specifically interested in the individual. Thus, where functional brain imaging is proposed for the detection of deception, it must be clear that the imaging strategy to be employed will provide satisfactory imaging data for valid interpretation (i.e., images of high statistical quality).[1]

LESSONS FROM THE POLYGRAPH

In 2003, the National Academy of Sciences (NAS) made a series of recommendations in its report on *The Polygraph and Lie Detection*. Although theses recommendations were primarily spawned by a consideration of the polygraph, they are relevant to the issues raised by the use of functional brain imaging as a tool for the detection of deception.[2]

Most people think of specific incidents or crimes when they think of lie detection. For example, an act of espionage has been committed and a suspect has been arrested. Under these circumstances the polygraph seems to perform above chance. The reason for this, the NAS committee believed, was something that psychologists have called the "bogus pipeline": If a person sincerely believed a given object (say a chair attached to electrical equipment) was a lie detector and that person was wired to the object and had committed a crime, a high probability exists (much greater than chance) that under interrogation the person would confess to the crime. The confession would have nothing to do with the basic scientific validity of the technique (i.e., the chair attached to electrical equipment) and everything to do with the individual's belief in the capability of the device to detect a lie. However, contrary to the belief that lie detection techniques such as the polygraph are most commonly used to detect the lies of the accused, by far the most important use of these techniques in the United States is in employee screening, pre-employment, and retention in high-security environments. The U.S. government performs tens of thousands of such studies each year in its various security

1. For a more in-depth explanation of functional brain imaging, see Raichle (2000); and Raichle and Mintun (2006).

2. I was a member of the NAS committee that authored the report.

agencies and secret national laboratories. This is a sobering fact given the concerns raised by the NAS report about the use of the polygraph in screening. As a screening technique the polygraph performs poorly and would likely falsely incriminate many innocent employees while missing the small number of spies in their midst. The NAS committee could find no available and properly tested substitute, including functional brain imaging, that could replace the polygraph.

The NAS committee found many problems with the scientific data it reviewed. The scientific evidence on means of lie detection was of poor quality with a lack of realism, and studies were poorly controlled, with few tests of validity. For example, the changes monitored (e.g., changes in skin conductance, respiration, and heart rate) were not specific to deception. To compound the problem, studies often lacked a theory relating the monitored responses to the detection of truthfulness. Changes in cardiac output, peripheral vascular resistance, and other measures of autonomic function were conspicuous by their absence. Claims with regard to functional brain imaging hinged for the most part on dubious extrapolations from group averages.

Countermeasures (i.e., strategies employed by a subject to "beat the polygraph") remain a subject clouded in secrecy within the intelligence community. Yet information on such measures is freely available on the Internet! Regardless, countermeasures remain a challenge for many techniques, although one might hold some hope that imaging could have a unique role here. For example, any covert voluntary motor or cognitive activity employed by a subject would undoubtedly be associated with predictable changes in functional brain imaging signals.

At present we have no good ways of detecting deception despite our very great need for them. We should proceed in acquiring such techniques and tools in a manner that will avoid the problems that have plagued the detection of deception since the beginning of recorded history. Expanded research should be administered by organizations with no operational responsibility for detecting deception. This research should operate under normal rules of scientific research with freedom and openness of communication to the extent possible while protecting national security. Finally, the research should vigorously explore alternatives to the polygraph, including functional brain imaging.

REFERENCES

National Academy of Sciences. 2003. *The polygraph and lie detection.* Washington, DC: National Research Council.

Raichle, M. E. 2008. A brief history of human brain mapping. *Trends in Neuroscience* (S0166-2236(08)00265-8[pii]10.1016/j.tins.2008.11.001).

Raichle, M. 2000. A brief history of human functional brain mapping. In *Brain mapping: The systems,* ed. A. Toga and J. Mazziotta. San Diego: Academic Press.

Raichle, M. E., and M. A. Mintun. 2006. Brain work and brain imaging. *Annual Review of Neuroscience* 29:449–476.

CHAPTER 2

The Use of fMRI in Lie Detection: What Has Been Shown and What Has Not

NANCY KANWISHER

Can you tell what somebody is thinking just by looking at magnetic resonance imaging (MRI) data from their brain?[1] My colleagues and I have shown that a part of the brain we call the "fusiform face area" is most active when a person looks at faces (Kanwisher et al. 1997). A separate part of the brain is most active when a person looks at images of places (Epstein and Kanwisher 1998). People can selectively activate these regions during mental imagery. If a subject closes her eyes while in an MRI scanner and vividly imagines a group of faces, she turns on the fusiform face area. If the same subject vividly imagines a group of places, she turns on the place area. When my colleagues and I first got these results, we wondered how far we could push them. Could we tell just by looking at the fMRI data what someone was thinking? We decided to run an experiment to determine whether we could tell in a single trial whether a subject was imagining a face or a place (O'Craven and Kanwisher 2000).

My collaborator Kathy O'Craven scanned the subjects, and once every twelve seconds said the name of a famous person or a familiar place. The subject was instructed to form a vivid mental image of that person or place. After twelve seconds Kathy would say, in random order, the name of another person or place. She then gave me the fMRI data from each subject's face and place areas.

Figure 1 shows the data from one subject. The x-axis shows time, and the y-axis shows the magnitude of response in the face area (black) and the place area (gray). The arrows indicate the times at which instructions were given to the subject. My job was to look at these data and determine for each trial whether the subject was imagining a face or a place. Just by eyeballing the data, I correctly determined in over 80 percent of the trials whether the subject was imagining faces or places. I worried for a long time before we published these data that people might think we could use an MRI to read their minds. Would they not realize the results obtained in my experiment were for

1. This article is based on remarks made at the American Academy of Arts and Sciences's conference on February 2, 2007.

a specific, constrained situation? That we used faces and places because we know which highly specific parts of the brain process those two categories? That we selected only cooperative subjects who were good mental imagers? And so on. I thought, "Surely, no one would try to use fMRI to figure out what somebody else was thinking!"

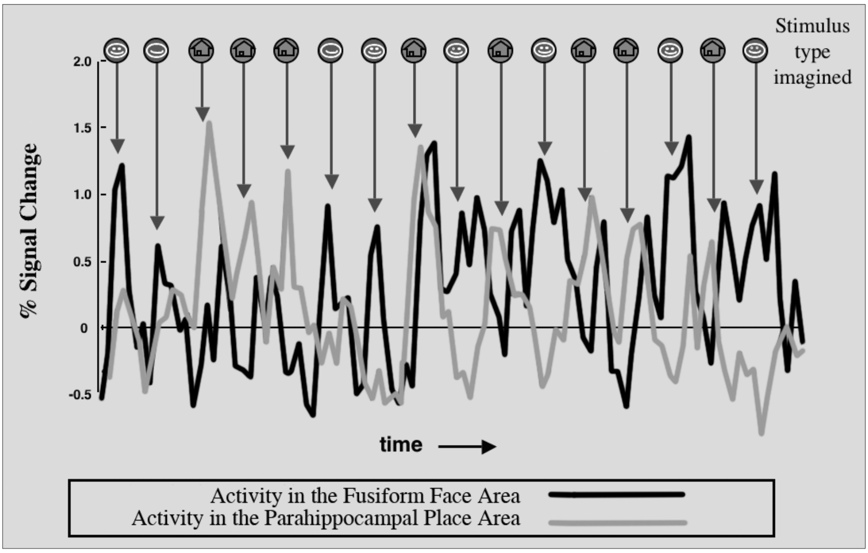


Figure 1. Time course of fMRI response in the fusiform face area and parahippocampal place area of one subject over a segment of a single scan, showing the fMRI correlates of single unaveraged mental events. Each black arrow indicates a single trial in which the subject was asked to imagine a specific person (indicated by face icon) or place (indicated by house icon). Visual inspection of the time courses allowed 83 percent correct determination of whether the subject was imagining a face or a place. Source: O'Craven and Kanwisher 2000.

My concern proved to be premature. Almost no one cited this work for several years. In the past couple of years, however, our findings have been more widely discussed—for example, in *Time Magazine*, on NPR, and in the *New York Times Magazine* (though oddly these venues still fail to cite our paper when describing our results)—and at least two companies, Cephos Corp. and No Lie MRI, have begun marketing fMRI "lie-detection" services. The Cephos website says, "Lying is shown to activate specific, discrete parts of the brain. We can use those regions to determine if a person is lying with a high degree of accuracy. No activation is seen when telling the truth." The No Lie MRI website includes a product overview that boasts, "Current accuracy is over 90% and is estimated to be 99% once product development is complete" (No Lie MRI 2006).

Does any real science lie behind these claims? The literature on using fMRI for lie detection can be divided into two groups (see Figure 2). The first uses group analysis, and the second focuses on individual subject data. In group analysis, subjects' brain scans are aligned as closely as possible. The data are then averaged and analyzed as a group. Group studies can be useful for identifying patterns of brain response that are consistent across subjects. Group studies are not useful for determining whether a particular subject is lying. Studies that analyze individual subject data *are* relevant for trying to determine whether fMRI is useful for lie detection, so we discuss these findings in turn.

I. Group Studies

- Ganis et al. 2003. *Cerebral Cortex* 13:830–836.
- Langleben et al. 2002. *Neuroimage* 15:727–732.
- Lee et al. 2002. *Human Brain Mapping* 15:157–164.
- Kozel et al. 2004. *Behavioral Neuroscience* 118:852–856.
- Kozel et al. 2004. *Journal of Neuropsychiatry and Clinical Neurosciences* 16:295–305.
- Nunez et al. 2005. *Neuroimage* 25:267–277.
- Mohmed et al. 2006. *Radiology* 238:679–688.

II. Individual Subject Studies

- Kozel et al. 2005. *Biological Psychiatry* 58:605–613.
- Davatzikos et al. 2005. *Neuroimage* 28:663–668.
- Langleben et al. 2005. *Human Brain Mapping* 26:262–272.

Figure 2. List of empirical studies of fMRI lie detection published in peer-reviewed journals as of January 2007.

STUDY 1

Kozel and colleagues (2005) instructed subjects to "steal" either a watch or a ring and to place the "stolen" item with their belongings in a locker while being watched by one of the investigators. They were then scanned with fMRI while answering a set of questions presented in a random order. For example, they might be asked, "Did you take the watch from the drawer?" Or, "Is the ring in your locker?" There were twenty roughly synonymous versions of each question, as well as a variety of "neutral" conditions; for example, "Do you like to swim?" Subjects were instructed to answer questions about the ring or watch as if they had stolen neither object and to answer the other questions truthfully. Thus, for the ring condition subjects were "lying" and for the watch condition they were "telling the truth," or vice versa. Participants were also told that they would receive $50 if the other investigator could not tell from the MRI data when they were lying.

What did they find? Well, first let me explain the basic logic of the analysis. Simply put, the idea is to take half of the subjects to discover some kind of brain "signature" of lying versus truth, and then to use that signature to classify the other subjects to figure out whether they were lying or telling the truth. More specifically, they found seven activation "blobs" in the brain that produced a stronger response during the lie condition than the truth condition, and they found the three or four of those blobs that were most discriminative between truths and lies. They then used those blobs to classify the activations in the other subjects. So, for example, they ran statistical tests on each 3-D pixel or "voxel" in the brain, asking whether that voxel produced a stronger response during the lie condition than the neutral condition, and they tallied how many voxels showed that pattern versus how many produced a stronger response in the truth condition than neutral. If they found more lie voxels than truth voxels, they considered their model to have identified which condition was a lie in that subject. By this measure, they could correctly determine for 90 percent of subjects which was the lie and which was the truth.

This is not really lie detection. The researchers always know that the subjects are lying in response to one of the sets of non-neutral questions. Rather than answering the question "Can you tell whether the subject is lying," this research is answering the question "Can you tell *which* response is the truth and which is the lie?"

STUDY 2

Langleben and colleagues (2005) scanned twenty-six subjects. Prior to the scan, a researcher presented the subjects with an envelope containing two playing cards—the seven of spades and the five of clubs—and a $20 bill. Participants were instructed to deny possession of one of the cards and acknowledge possession of the other card during the scans. They were also told they could keep the $20 if they successfully concealed the identity of the "lie" card. A different researcher then scanned the subjects, telling them to respond to each trial as accurately and truthfully as possible. During the scans the subjects saw, in random order, the five of clubs and seven of spades, and they responded with a button press, indicating whether they held that card in their possession. Of course, one of these card types would be the lie condition, and one would be the truth condition. Other various cards were included as control conditions. Critically, the truth condition was the only response for which the subjects said "yes," and there were only 24 of these yes-truth trials out of 432 total trials. This feature is important because it means that the subjects are sitting during the scanning saying, "No, no, no, no, no, no," most of the time, looking for those rare five of clubs so that they can say yes. The subjects probably think of the task essentially as one of detecting that five of clubs, and it means that the neural signature of the supposed "truth" response is really just the neural signature of a target detection event.

When the researchers analyzed their data using group analysis, they found no significantly higher response for the truth condition than for the lie condition, thus failing to replicate their own 2002 study. The probable reason for this failure is that the truth condition was the salient target detection condition, and so the lie condition was like a default case and hence did not activate any part of the brain to a greater extent. Next the researchers lowered the statistical threshold to $p < .05$ (uncorrected). However, the brain contains 30,000–40,000 voxels, so $p < .05$ creates a situation in which hundreds or thousands of voxels reach significance, even if what is being analyzed is random noise. Neuroimagers must correct for this "multiple comparisons" problem and generally do not accept the .05 threshold as legitimate without such corrections. Nonetheless, using $p < .05$ (uncorrected), Langleben and colleagues reported that they found activations in the left inferior frontal gyrus for lie versus truth, commenting that "Lie related activation was less robust and limited to areas associated with context processing, left inferior frontal gyrus" (Langleben et al. 2005). In fact, this "result" was not even worth this effort at interpretation, because it is nowhere near significant.

In addition to the group analysis, Langleben and colleagues also performed an individual subject analysis in which they asked whether they could discriminate lies from truths. They used a classification method based on the data from the group analysis to test new subjects. They found that they could discriminate, on an individual trial basis, which responses were lies and which were truths 76 percent of the time. The "false alarm" rate, which is the rate of true responses incorrectly labeled lies, was 16 percent, and the miss rate, which is the rate of false responses incorrectly labeled truths, was 31 percent. However, because in this experiment the truths were rare target events, the ability to discriminate truths from lies probably just reflects the ability to distinguish rare events; it has nothing to do with lying per se.

In the final paper of the three that look at individual subjects, Davatzikos et al. (2005) analyzed the same data in the Langleben et al. paper (2005). They just used fancier math. There is a lot of exciting work going on right now in computer science and math, where people are devising machine learning algorithms to find patterns in MRI data and other kinds of neuroscience data. So they used some of these fancier methods to classify the responses, and they got an 89 percent correct classification on a subject basis, not a trial basis. But now we have to consider what this means, and whether these lab experiments have anything to do with lie detection as it might be attempted in the real world.

REAL-WORLD IMPLICATIONS

To summarize all three of the individual subject studies, two sets of functional MRI data have been analyzed and used to distinguish lies from truth. Kozel and colleagues (2005) achieved a 90 percent correct response rate in determining which was the lie and which was the truth, when they knew in advance there would be one of each. Langleben got a 76 percent correct response rate with individual trials, and Davatzikos, analyzing the same data, got an 89 percent correct response rate. The very important caveat is that in the last two studies it is not really lies they were looking at, but rather target detection events. Leaving that problem aside, these numbers aren't terrible. And these classification methods are getting better rapidly. Imaging methods are also getting better rapidly. So who knows where all this will be in a few years. It could get even much better than that.

But there is a much more fundamental question. What does any of this have to do with real-world lie detection? Let's consider how lie detection in the lab differs from any situation where you might want to use these methods in the real world. The first thing I want to point out is that making a false response when you are instructed to do so isn't a lie, and it's not deception. It's simply doing what you are told. We could call it an "instructed falsehood." Second, the kind of situation where you can imagine wanting to use MRI for lie detection differs in many respects from the lab paradigms that have been used in the published studies. For one thing, the stakes are incomparably higher. We are not talking about $20 or $50; we are talking about prison, or life, or life in prison. Further, the subject is suspected of a very serious crime, and they believe while they are being scanned that the scan may determine the outcome of their trial. All of this should be expected to produce extreme anxiety. Importantly, *it should be expected to produce extreme anxiety whether the subject is guilty or not guilty of the crime.* The anxiety does not result from guilt per se, but rather simply from being a suspect. Further, importantly, the subject may not be interested in cooperating, and all of these methods we have been discussing are completely foilable by straightforward countermeasures.

Functional MRI data are useless if the subject is moving more than a few millimeters. Even when we have cooperative subjects trying their best to help us and give us good data, we still throw out one of every five, maybe ten, subjects because they move too much. If they're not motivated to hold still, it will be much worse. This is not just a matter of moving your head—you can completely mess up the imaging data just by moving your tongue in your mouth, or by closing your eyes and not being able to read the questions. Of course, these things will be detectable, so the experimenter would know that the subject was using countermeasures. But there are also countermeasures subjects could use that would not be detectable, like performing mental arithmetic. You can probably activate all of those putative lie regions just by subtracting seven iteratively in your head.

Because the published results are based on paradigms that share none of the properties of real-world lie detection, those data offer no compelling evidence that fMRI will work for lie detection in the real world. No published evidence shows lie detection with fMRI under anything even remotely resembling a real-world situation. Furthermore, it is not obvious how the use of MRI in lie detection could even be tested under anything resembling a real-world situation. Researchers would need access to a population of subjects accused of serious crimes, including, crucially, some who actually perpetrated the crimes of which they are accused and some who did not. Being suspected but innocent might look a lot like being suspected and guilty in the brain. For a serious test of lie detection, the subject would have to believe the scan data could be used in her case. For the data from individual scans to be of any use in testing the method, the experimenter would ultimately have to know whether the subject of the scan was lying. Finally, the subjects would have to be interested in cooperating. Could such a study ever be ethically conducted?

Before the use of fMRI lie detection can be seriously considered, it must be demonstrated to work in something more like a real-world situation, and those data must be published in peer-reviewed journals and replicated by labs without a financial conflict of interest.

REFERENCES

Davatzikos, C., K. Ruparel, Y. Fan, D. G. Shen, M. Acharyya, J. W. Loughead, R. C. Gur, and D. D. Langleben. 2005. Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *Neuroimage* 28:663–668.

Epstein, R., and N. Kanwisher. 1998. A cortical representation of the local visual environment. *Nature* 392:598–601.

Kanwisher, N., J. McDermott, and M. Chun. 1997. The fusiform face area: A module in human extrastriate cortex specialized for the perception of faces. *Journal of Neuroscience* 17:4302–4311.

Kozel, F., K. Johnson, Q. Mu, E. Grenesko, S. Laken, and M. George. 2005. Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry* 58:605–613.

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

No Lie MRI. 2006. Product overview. http://www.noliemri.com/products/Overview.htm.

O’Craven, K., and N. Kanwisher. 2000. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience* 12:1013–1023.

CHAPTER 3

Lying Outside the Laboratory: The Impact of Imagery and Emotion on the Neural Circuitry of Lie Detection

ELIZABETH A. PHELPS

One of the challenges of research on lie detection is the difference between instructed lying in a laboratory setting and the types of situations one might encounter outside the laboratory that would require the use of lie detection techniques. The development of techniques for lie detection is based on laboratory studies of lying. However, the characteristics of lies outside the laboratory may differ in important ways. For instance, if someone is accused of a crime, there are two possible scenarios. First, the person may be innocent. If this is the case, he or she is likely upset and wondering how the accusers could think the charges are plausible. Given the serious circumstances, this person might ruminate on this last point and be quite concerned about proving his or her innocence. On the other hand, if a person is accused of a crime he or she actually committed, there might be an anxious or guilty feeling and an effort to formulate a false alibi. This person might think about this lie in detail and elaborate on the lie in an effort to be especially convincing when asked about the alibi. In both cases, the person accused of the crime faces a highly emotional situation and has ample opportunity to mentally image the circumstances that occurred and ruminate on the situation. It is these factors, *emotion* and *imagery,* that differ between the real lie and the laboratory lie. Research in cognitive neuroscience has shown that both imagery and emotion can alter the representation of events. Given that lies outside the laboratory are likely to be personally relevant and emotional, and also imagined with elaboration and repetition, any successful lie detection techniques will need to consider these factors. In this essay, I will explore some of the ways imagery and emotion might impact the potential neural signatures of lying.

There are two possible uses of functional magnetic resonance imaging (fMRI) for lie detection. One is to assess familiarity. Imagine that the police want to know whether a suspect has been to the scene of a crime. The suspect might deny any familiarity with that location. In using fMRI to assess familiarity, the police could show the suspect a picture of the scene and look for a pattern of blood oxygenation level dependent (BOLD) signal in his or her brain that indicates previous experience with that scene. In this case, the use of fMRI for lie detection is based on our knowledge about the neural representation of *memory*. The second use of fMRI for lie detection is to detect deception. One would expect that the neural systems involved when someone is lying or telling the truth are different. The assumption underlying lie detection techniques for detecting deception is that the truth is the natural response. When individuals lie they have to inhibit the truth to generate the lie. In this case, lying results in conflict between the truth and the lie. The use of fMRI to detect deception is based on our knowledge about the neural representation of *conflict*. When lying outside the laboratory, the stakes are high, the individual is highly emotional, and the story is practiced, rehearsed, or imagined. Because of this, the use of fMRI for lie detection will need to consider how imagery and emotion might alter the neural representation of memory and conflict.

To address these questions, I will first review what is known about the neural signatures of using fMRI to detect familiarity. Have we identified reliable neural markers for item or event familiarity? In other words, when presented with a face or scene seen before or that is reminiscent, does the brain respond with a pattern of activity that is different in ways that can be measured with fMRI? This question was addressed in a recent study by Gonsalves, Wagner, and colleagues (2005). In this study, the participants were presented with a series of faces. Afterwards, they were given a recognition test in which some of the faces were presented earlier, some were morphed to look somewhat like the faces presented earlier, and others were novel. For each face, participants were asked to judge whether they recollected having seen the face earlier, if it seemed familiar, or if it was new. There were a few regions in the temporal lobe, the hippocampal cortex and fusiform gyrus, where BOLD responses differed depending on the mnemonic judgment. The term hippocampal cortex refers to a collection of regions known to be important for memory, including the hippocampus proper and regions around and underneath it, such as the parahippocampal cortex. These regions showed more activity when the faces were judged to be less familiar than when they were more familiar. The fusiform gyrus, which plays an important role in processing faces, showed more activity when faces were more familiar. These results suggest the possibility of a neural signature for familiarity that could be detected with fMRI (see Figure 1).

However, what do we know about how responses in the hippocampus, parahippocampus, and fusiforn gyrus might be altered with imagery and emotion? First, let's explore the role of imagery. A classic paradigm demonstrates the importance of imagery in memory. Imagine you were presented the following list of words: SOUR, CANDY, SUGAR, BITTER, GOOD,

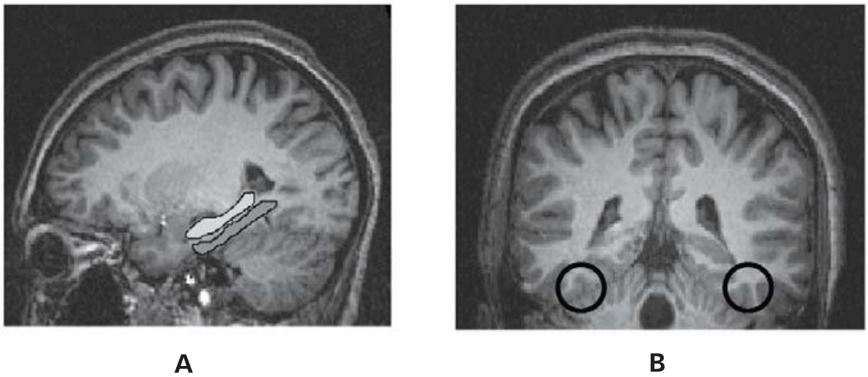


Figure 1. (A) the hippocampus (light grey) and parahippocampus (dark grey); (B) the fusiform face area (in circles)

TASTE, TOOTH, NICE, HONEY, SODA, CHOCOLATE, HEART, CAKE, TART, PIE. After the presentation of this list and a short delay you are given a recognition test. For instance, you might be asked if *tart* was on the list, in which case you would say "yes." If you were asked if *chair* was on the list, you would correctly respond "no." However, what if you were asked if *sweet* was on the list? If you go back a few sentences, you will see that *sweet* was not on the list, however most participants in experiments like this say "yes" it was on the original list. This is because sweet is strongly associated with all the words on the list. Even though *sweet* was not on the list, your mind most likely came up with the image of sweet when you were reading the words. This type of mistake is often called a false memory, but is not really a *false* memory. We have memories both for images that we internally generate and events that occur in the external world. In this case, the word *sweet* is really an imagined memory, something that was generated when you saw the initial list of words, and therefore, most people misremember actually having seen it before. This is an example of mental imagery creating a memory that results in a memory mistake.

Can fMRI images of the brain distinguish imagined or false memories from memories based on perceptual experience? As the study by Gonsalves and colleagues (2005) demonstrates, fMRI signals can indicate familiarity, but what happens when an item is familiar because the individual has imagined it? Using the word list paradigm described above it was shown that some regions known to be important for memory, such as the hippocampus, do not differentiate memories for events that are perceptually experienced from events that are imagined (Cabeza et al., 2001). In other words, our memories for real and imagined events rely on overlapping neural systems. If we were to look at responses in the hippocampus, we could not differentiate if an event is familiar due to perceptual experience or mental imagery. However, there are other regions of the hippocampal cortex, specifically the parahippocampus, that are more important in processing the perceptual details of memory. In this same study, this region showed greater activation to true relative to false memories. This study, and others like it (see Schacter and

Slotnick, 2004 for a review), indicates that for many brain regions known to be important in memory, such as the hippocampus, it does not matter whether an experienced event was the result of our thought or our perception. Given this, we cannot look at these memory-related brain regions to reveal whether a person is remembering accurately. Other brain regions, such as the parahippocampus and fusiform gyrus, are more specifically involved in perceptual processing and memory for perceptual details (Schacter and Slotnick, 2004). These regions may provide a signal of familiarity for scenes and faces.

However, one difficulty in relying on BOLD signal responses in regions involved in perceptual aspects of memory, such as the parahippocampus or fusiform gyrus, to judge whether a suspected criminal is familiar with a scene or face is that perception is often altered by emotion. For the criminal whose brain is being imaged to judge involvement in a crime, pictures of the scene of the crime or partners in crime are likely highly emotional. Changes in perception occur with emotion, and research has demonstrated changes in BOLD signal in both the parahippocampus and fusiform gyrus for emotional scenes and faces. For example, a study looking at individuals remembering 9/11 found that people who were closer to the World Trade Center showed less activity in the parahippocampus when recalling the events of 9/11 than when recalling less emotional events (Sharot et al., 2007). As indicated earlier, the parahippocampus also shows less activation when a face is more familiar (Gonsalves et al., 2005). Even though imaging this region might reveal perceptual qualities of memory, this region might also be influenced by emotion. If the event is highly emotional, signals in the parahippocampus might not be a reliable indicator of familiarity.

What about the fusiform gyrus? This region is known for processing faces (Kanwisher and Yovel, 2006, for a review). As indicated by Gonslaves and colleagues (2005) this region also shows stronger activation for more familiar faces. However, emotion also influences responses in the fusiform gyrus, so that in a highly emotional situation the signal from this region might be somewhat altered. For faces that are equally unfamiliar, more activation is observed in the fusiform gyrus for faces with fear expressions (Vuilleumier et al., 2001). Furthermore, the face itself does not need to be fearful. If the context in which the face is presented is fearful, greater activation of the fusiform gyrus is observed (Kim et al., 2004). Because of this, responses in this region may not be a reliable indicator of familiarity with a face in an emotional situation.

When researchers look at the brain's memory circuitry to detect familiarity with a scene or person, for some regions it may be difficult to differentiate events that a person imagined, rehearsed, or thought were plausible from those that actually occurred. Memories are formed for events that happen only in our minds and events that happen in the outside world. The regions that are important in the perceptual aspects of memory are influenced by emotion, so they might not be good detectors of familiarity if the situation is emotional. In other words, the imagery and emotion that would likely be present when a lie is personally relevant and important might interfere with the use of fMRI signals to detect familiarity.

The second potential use of fMRI for lie detection relies on our knowledge of the neural circuitry of conflict. It is assumed that lying results in a conflict between the truth and the lie. How might emotion and imagery influence the neural signatures of conflict? Two regions that have been highlighted for their role in responding to conflict or interference are the anterior cingulate cortex and the inferior frontal gyrus. These regions have also been implicated in studies of lie detection (e.g., Kozel et al., 2004; Langelben et al., 2005). One classic task used to detect conflict-related responses is the Stroop test, in which participants are shown a list of words and asked not to read the words but to name the colors in which the words are printed. For instance, if the word *table* is presented in blue ink, participants are asked to say "blue." Most of the time, participants can fairly easily ignore the words and name the color of the ink. However, if the words the participants are asked to ignore are the names of colors, it is much more difficult. For example, it typically takes significantly longer for participants to name the ink color "blue" if the word they are asked to ignore is *red* as opposed to *table*. This longer reaction time is due to the conflict between reading the word *red* and saying the word "blue." Naming the color of ink for color words in comparison to other words results in significant activation of an anterior, dorsal region of the cingulate cortex (Carter and van Veen, 2007) and this same region shows activation in many laboratory studies of lie detection (e.g., Kozel et al., 2004).

However, difficulty in naming the ink color of words is not only slower for color words. In a variation of the Stroop task, called the emotional Stroop task, subjects are presented with highly emotional words printed in different colors of ink. When asked to ignore the words and name the ink color, it takes significantly longer to name the color for emotional words in comparison to neutral words. Interestingly, emotional variations of the Stroop task also result in activation of the anterior cingulate, but a slightly different region that is more ventral than that observed in the classic Stroop paradigm (Whalen et al., 1998). In a meta-analysis of a number of conflict tasks, Bush et al. (2000) confirmed this division within the anterior cingulate (see Figure 2). Cognitive conflict tasks, such as the classic Stroop task, typically result in activation of the dorsal anterior cingulate, whereas emotional conflict tasks, as demonstrated by the emotional Stroop task, result in activation of the ventral anterior cingulate. This suggests that this specific neural indicator of conflict is significantly altered depending on the emotional nature of the conflict.

Another region often implicated in conflict or interference in studies of lie detection is the inferior frontal gyrus. In fact, some studies of lie detection have suggested that activation of this region is the best predictor of whether a participant is lying (Langleben et al., 2005). The role this region plays in conflict or interference monitoring has traditionally been examined with the Sternberg Proactive Interference paradigm. In a typical version of this paradigm, a participant is shown a set of stimuli and told to remember it. For example, the set might include three letters, such as B, D, F. After a short delay the participant is presented a letter and asked, "Was this letter in the target set?" If the letter is *D*, the participant should answer "yes." In the next

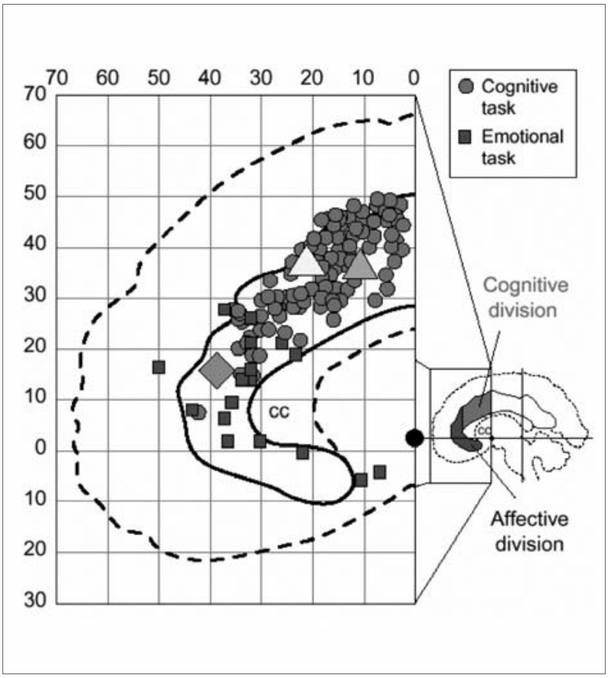


Figure 2. Meta-analysis of fMRI studies showing anterior cingulate activation for cognitive (circles) and emotional (squares) tasks demonstrating a cognitive-affective division within the anterior cingulate. Consistent with this division, cognitive (triangles) and affective (diamond) versions of a Stroop task results in activation of dorsal and ventral regions of the anterior cingulate, respectively. Reprinted with permission from Bush et al., 2000.

trial the participant is given another target set, such as K, E, H. At this point, if the participant is shown the letter *P*, she or he should say "no." If the participant is shown the letter *B*, the correct answer is also "no." However, for most participants it will take longer to correctly respond "no" to *B* than *P*. This is because *B* was a member of the immediately preceding target set (B, D, F), but it is not a member of the current target set (K, E, H). On the preceding trial, a minute or so earlier, the participant was ready to respond "yes" to *B*. To correctly respond "no" to *B* on the current trial requires the participant to inhibit this potential "yes" response and focus only on the current target set. This requirement for inhibition is not necessary if the probe letter is *P*, which was not a member of either the preceding or current target set. Research using both brain imaging (D'Esposito et al., 1999; Jonides and Nee, 2006) and lesion (Thompson-Schill et al., 2002) techniques has shown that the inferior frontal gyrus plays an important role in resolving this type of interference or conflict. It is believed this region might be linked to lying because in order to lie one must inhibit the truth, which creates conflict or interference (e.g., Langleben et al., 2005).

In order to examine the impact of emotion on this type of interference, a recent study used a variation of the typical Sternberg Proactive Interference paradigm in which the stimuli were emotional words or scenes, instead of letters or neutral words and scenes. An examination of reaction times found that the inhibition of emotional stimuli was faster than neutral stimuli, suggesting that emotion can impact processing in this interference paradigm (Levens and Phelps, 2008). Using fMRI to examine the neural circuitry underlying the impact of emotion on the Sternberg Proactive Interference paradigm revealed that the inhibition of emotional stimuli on this task engages a slightly different network, including regions of the anterior insula cortex and orbitofrontal cortex (Levens et al., 2006). Much like the results observed with the anterior cingulate, this suggests that emotion alters the neural circuitry of inhibition or conflict observed in the inferior frontal gyrus. Although further studies are needed to clarify the impact of emotion on interference resolution mediated by the inferior frontal gyrus, these initial results suggest that this neural indicator of conflict in lying may also be different in highly emotional situations.

There is abundant evidence that emotion can influence the neural circuitry of conflict or interference identified in laboratory studies of lie detection, but can imagery or repetition also alter responses in these regions? It seems possible that practicing a lie repeatedly, as one might after generating a false alibi, could reduce the conflict experienced when telling that lie. If this is the case, we might expect less evidence of conflict with practice or repetition. This finding has been observed with the classic Stroop paradigm. A number of behavioral studies have demonstrated that practice can diminish the Stroop effect (see MacLeod, 1991, for a review). It has also been shown that practicing the Stroop task significantly reduces conflict-related activation in the anterior cingulate (Milhan et al., 2003). To date, there is little research examining how imagery might alter the neural circuitry of conflict or interference as represented in the inferior frontal gyrus, but these findings suggest that at least some neural indicators of conflict or interference may be unreliable if the task (or lie) is imagined, practiced, and rehearsed.

Although the use of fMRI to detect lying in legal settings holds some promise, there are some specific challenges in developing these techniques that have yet to be addressed. Out of necessity, the development of techniques for lie detection relies on controlled laboratory studies of lying. However, lying in legally relevant circumstances is rarely so controlled. This difference should be kept in mind when building a neurocircuitry of lie detection that is based on unimportant lies told by paid participants in the laboratory, but is intended to be applied in legally important situations to people outside the laboratory facing far higher stakes. Because of this, it is important to examine exactly what might differ between the laboratory lie and the other lies that could impact the usefulness of these techniques. In this essay, I have explored two factors–imagery and emotion–and highlighted how research suggests that the neural signatures identified in current fMRI lie detection technologies might be quite different in their utility when the lies detected are not generated in the laboratory. This problem of applying laboratory findings to other, more everyday and/or personally relevant and important circumstances is a challenge for all studies of human behavior. However, addressing this challenge becomes especially critical when we attempt to use our laboratory findings to generate techniques that can potentially impact individuals' legal rights. Until this challenge can be addressed, the use of fMRI for lie detection should remain a research topic, instead of a legal tool.

Bush, G., P. Luu, and M.I. Posner. 2000. Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Science* 4:215–222.

Cabeza, R., S.M. Rao, A.D. Wagner, A.R. Mayer, and D.L. Schacter. 2001. Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences* 98:4805–4810.

Carter, C.S., and V. van Veen. 2007. Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, Behavioral Neuroscience* 7:367–379.

D'Esposito, M., B.R. Postle, J. Jonides, and E.E. Smith. 1999. The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. *Proceedings of the National Academy of Sciences* 96:7514–7519.

Gonsalves, B.D., I. Kahn, T. Curran, K.A. Norman, and A.D. Wagner. 2005. Memory strength and repetition suppression: Multimodal imaging and medial temporal cortical contributions to recognition. *Neuron* 47:751–781.

Jonides, J., and D.E. Nee. 2006. Brain mechanisms of proactive interference in working memory. *Neuroscience* 139:181–193.

Kanwisher, N., and G. Yovel. 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society: B Biological Science* 361:2109–2128.

Kim, H., L.H. Somerville, T. Johnstone, S. Polis, A.L. Alexander, L.M. Shin, and P.J. Whalen. 2004. Contextual modulation of amygdala responsivity to surprised faces. *Journal of Cognitive Neuroscience* 16:1730–1745.

Kozel, F.A., T.M. Padgett, and M.S. George. 2004. A replication study of the neural correlates of deception. *Behavioral Neuroscience* 118:852–856.

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

Levens, S.M., M. Saintilus, and E.A. Phelps. 2006. Prefrontal cortex mechanisms underlying the interaction between emotion and inhibitory processes in working memory. Cognitive Neuroscience Society, San Francisco, CA.

Levens, S.M., and E.A. Phelps. 2008. Emotion processing effects on interference resolution in working memory. *Emotion* 8:267–280.

MacLeod, C.M. 1991. Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin* 109:163–203.

Milhan, M.P., M.T. Banich, E.D. Claus, and N.J. Cohen. 2003. Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *Neuroimage* 18:483–493.

Schacter, D.L., and S.D. Slotnick. 2004. The cognitive neuroscience of memory distortion. *Neuron* 44:149–160.

Sharot, T., E.A. Martorella, M.R. Delgado, and E.A. Phelps. 2007. How personal experience modulates the neural circuitry of memories of September 11. *Proceedings of the National Academy of Sciences* 104:389–394.

Thompson-Schill, S.L., J. Jonides, C. Marshuetz, E.E. Smith, M. D'Esposito, I.P. Kan, R.T. Knight, and D. Swick. 2002. Effects of frontal lobe damage on interference effects in working memory. *Cognitive, Affective, Behavioral Neuroscience* 2:109–120.

Vuilleumier, P., J.L. Armony, J. Driver, and R.J. Dolan. 2001. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron* 30:829–841.

Whalen, P.J., G. Bush, R.J. McNally, S. Wilhelm, S.C. McInerney, M.A. Jenike, and S.L. Rauch. 1998. The emotional counting Stroop paradigm: a functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry* 44:1219–1228.

CHAPTER 4

Actions Speak Louder than Images[1]

STEPHEN J. MORSE

INTRODUCTION

Law must answer two types of general questions: 1) What legal rules should govern human interaction in a particular context? and 2) How should an individual case be decided? Scientific information, including findings from the new neurosciences, can be relevant both to policy choices and to individual adjudication. Most legal criteria are behavioral, however, including, broadly speaking, actions and mental states, and it could not be otherwise. The goal of law is to help guide and order the interactions between acting persons. Consider criminal responsibility, the legal issue to which neuroscience is considered most relevant. Criminal prohibitions all concern culpable actions or omissions and are addressed to potentially rational persons, not to brains. Brains do not commit crimes. Acting people do. We do not blame and punish brains. We blame and punish persons if they culpably violate a legal prohibition that society has enacted. All legally relevant evidence, whether addressed to a policy choice or to individual adjudication, must therefore concern behavior entirely or in large measure.

Behavioral evidence will almost always be more legally useful and probative than neuroscientific information. If no conflict exists between the two types of evidence, the neuroscience will be only cumulative and perhaps superfluous. If conflict does exist between behavioral and neuroscientific information, the strong presumption must be that the behavioral evidence trumps the neuroscience. Actions speak louder than images. If the behavioral evidence is unclear, however, but the neuroscience is valid and has legally relevant implications, then the neuroscience may tip the decision-making balance. The question is whether neuroscientific (or any other) evidence is legally relevant; that is, whether it genuinely and precisely helps answer a question the law asks.

Consider the following examples of both types of questions, beginning with a general legal rule. Should adolescents who culpably commit capital murder when they are sixteen or seventeen years old qualify for imposition

1. The title of this paper is a precise copy of the title of an article by Apoorva Mandavilli that appeared in *Nature* in 2006.

of the death penalty, or should that punishment be categorically barred for this class of murderers? Recent neuroscience evidence has demonstrated that the adolescent frontal cortex—roughly, the seat of judgment and behavioral control—is not fully biologically mature. What is the relevance of this information to deciding whether the death penalty should be imposed (see Roper v. Simmons, 2005)?

Now consider the following case of individual adjudication. A sixty-three-year-old businessman with no history of violence or other antisocial conduct has a harsh argument with his wife. During the course of the argument, she lunges at him and scratches his face. He grabs her, strangles her to death, and then throws her out the window of their twelfth-story apartment. The husband is charged with murder. He has the means to pay for a complete psychiatric and neurological workup that discloses that he has a sizable but benign subarachnoid cyst pressing on his frontal cortex. What is the relevance of this finding to his culpability for homicide (see People v. Weinstein, 1992)?

FALSE STARTS

Some common misconceptions bedevil clear thinking about the relevance of neuroscience to law: the belief that scientific discoveries necessitate particular political or legal rules or institutions; the belief that neuroscientific explanation of behavior or determinism generally poses a threat to the legal concept of the person; and the belief that discovery of a cause for behavior means that within our current responsibility practices the agent is not responsible for the behavior, an error I have previously termed the "fundamental psycholegal error" (Morse 1994).

Politics, morality, and law all concern how human beings should live together. They are the domains of practical reason and normativity. As a discipline of theoretical reason, science can help us understand the causal variables that constrain and shape human behavior. Such information can and should inform our reasoning about how to live together, but it cannot dictate any particular answer because how we should live is not a matter of theoretical reason. Some moral theorists believe that we can deduce moral conclusions from facts about the world, but this position is entirely controversial and those who hold it often disagree about specific moral rules. Many people believe that the new neuroscience suggests that a fully physical explanation of all human behavior is possible and that human behavior is as determined as all the rest of the phenomena of the universe. Some conclude in response that we should adopt a fully consequential morality in which concepts like just deserts have no justifiable place, but this conclusion does not ineluctably follow from the truth of universal causation. Even if it did, it would not tell us which goods to maximize, nor would it provide a source of normativity.

To take a more specific example, a recent, provocative study showed that a particular type of brain damage was associated with a willingness directly to take innocent life to achieve a net saving of lives (Koenigs et al. 2007). People without such brain damage were willing indirectly to cause the death of an innocent person to save more lives. If they had to kill the victim directly, however, they stopped calculating and refused to take an innocent life even to save net lives. If the study is valid in real-world conditions, it suggests that people with "normal" brains do not consequentially calculate under some conditions and perhaps it suggests that socializing them to do so might be difficult. But this finding does not necessarily mean that people cannot be socialized to calculate if we thought that such consequential calculation was desirable.

The new neuroscience joins a long list of contenders for a fully causal, scientific explanation of human behavior, ranging from sociological to psychological to biological theories. Such explanations are thought to be threats to the law's conception of the person and responsibility. Neuroscience concerns the brain—the biological source of our humanity, personhood, and sense of self—and it seems to render the challenge to the legal concept of the person more credible. The challenge arises in two forms. The first does not deny that we are the types of creatures we think we are, but it simply assumes that responsibility and all that it implies are impossible if determinism is true. This is a familiar claim. The second challenge denies that we are the type of creatures we think we are and that is presupposed by law and morality. This is a new and potentially radical claim. Neither succeeds at present, however.

The dispute about whether responsibility is possible in a deterministic world has been ongoing for millennia, and no resolution is in sight (Kane 2005). No uncontroversial definition of determinism has been advanced, and we will never be able to confirm that it is true or not. As a working definition, however, let us assume, roughly, that all events have causes that operate according to the physical laws of the universe and that they were themselves caused by those same laws operating on prior states of the universe in a continuous thread of causation going back to the first state. Even if this is too strong, the universe seems so sufficiently regular and lawful that rationality demands that we must adopt the hypothesis that universal causation is approximately correct. The English philosopher, Galen Strawson, calls this the "reality constraint" (Strawson 1989). If determinism is true, the people we are and the actions we perform have been caused by a chain of causation over which we mostly had no rational control and for which we could not possibly be responsible. We do not have contra-causal freedom. How can responsibility be possible for action or for anything else in such a universe? How can it be rational and fair for civil and criminal law to hold anyone accountable for anything, including blaming and punishing people because they allegedly deserve to be blamed and punished?

Three common positions are taken in response to this conundrum: metaphysical libertarianism, hard determinism, and compatibilism. Libertarians believe that human beings possess a unique kind of freedom of will and action according to which they are "agent originators" or have "contra-causal freedom." In short, they are not determined and effectively able to act uncaused by anything other than themselves (although they are of course influenced by their time and place and can only act on opportunities that exist then). The buck stops with them. Many people believe that libertarianism is a foundational assumption for law. They believe that responsibility is possible only if we genuinely possess contra-causal freedom. Thus, if we do not have this extraordinary capacity, they fear that many legal doctrines and practices, especially those relating to responsibility, may be entirely incoherent. Nonetheless, metaphysical libertarianism is not a necessary support for current responsibility doctrines and practices. All doctrines of criminal and civil law are fully consistent with the truth of determinism (Morse 2007). Moreover, only a small number of philosophers and scientists believe that human beings possess libertarian freedom of action and will, which has been termed a "panicky" metaphysics (Strawson 1982) because it is so implausible (Bok 1998).

Hard determinists believe that determinism is true and is incompatible with responsibility. Compatibilists also believe that determinism is true but claim that it is compatible with responsibility. For either type of determinist, biological causes, including those arising from the brain, pose no new or more powerful general metaphysical challenge to responsibility than nonbiological or social causes. As a conceptual and empirical matter, we do not necessarily have more control over psychological or social causal variables than over biological causal variables. More important, in a world of universal causation or determinism, biological causation creates no greater threat to our life hopes than psychological or social causation. For purposes of the metaphysical free-will debate, a cause is just a cause, whether it is neurological, genetic, psychological, sociological, or astrological. Neuroscience is simply the newest "bogey" in a dispute about the general possibility of responsibility that has been ongoing for millennia. It certainly is more scientifically respectable than earlier bogeys, such as astrology and psychoanalysis, and it certainly produces compelling representations of the brain (although these graphics are almost always misleading to those who do not understand how they are constructed). But neuroscience evidence for causation does no more work in the general free-will/responsibility debate than other kinds of causal evidence.

Hard determinism does not try either to explain or to justify our responsibility concepts and practices; it simply assumes that genuine responsibility is metaphysically unjustified. For example, a central hard determinist argument is that people can be responsible only if they could have acted otherwise than they did, and if determinism is true, they could not have acted other than they did (Wallace 1994). Consequently, the hard determinist claims that even if an internally coherent account of responsibility and related practices can be given, it will be a superficial basis for responsibility, which is allegedly only an illusion (Smilansky 2000). There is no "real" or "ultimate" responsibility. Hard determinists concede that Western systems of law and morality hold some people accountable and excuse others, but the hard determinist argues that these systems have no justifiable basis for distinguishing genuinely responsible from nonresponsible people. Hard determinists sometimes accept responsibility ascriptions because doing so may have good consequences, but they still deny that people are genuinely responsible and robustly deserve praise and blame and reward and punishment.

Hard determinism thus provides an external critique of responsibility. If determinism is true and is genuinely inconsistent with responsibility, then no one can ever be really responsible for anything and desert-based responsibility attributions cannot properly justify further action. The question, then, is whether as rational agents we must swallow our pride, accept hard determinism because it is so self-evidently true, and somehow transform the legal system and our moral practices accordingly.

Compatibilists, who agree with hard determinists that determinism is true, have three basic answers to the incompatibilist challenge. First, they claim that responsibility attributions and related practices are human activities constructed by us for good reason and that they need not conform to any ultimate metaphysical facts about genuine or "ultimate" responsibility. Indeed, some compatibilists deny that conforming to ultimate metaphysical facts is even a coherent goal in this context. Second, compatibilism holds that our positive doctrines of responsibility are fully consistent with determinism. Third, compatibilists believe that our responsibility doctrines and practices are normatively desirable and consistent with moral, legal, and political theories that we firmly embrace. The first claim is theoretical; the third is primarily normative. Powerful arguments have been advanced for the first and third claims (Lenman 2006; Morse 2004). For the present purpose, however, which is addressed to whether free will is really foundational for law, the second claim is the most important.

The capacity for rationality is the primary responsibility criterion, and its lack is the primary excusing condition. Human beings have different capacities for rationality in general and in specific contexts. For example, young children in general have less developed rational capacity than adults. Rationality differences also differentially affect agents' capacity to grasp and to be guided by good reason. Differences in rational capacity and its effects are real even if determinism is true. Compulsion is also an excusing condition, but it is simply true that some people act in response to external or internal hard choice threats to which persons of reasonable firmness might yield, and most people most of the time are not in such situations when they act. This is true even if determinism is true and even if people could not have acted otherwise.

Consider the doctrines of criminal responsibility. Assume that the defendant has caused a prohibited harm. Prima facie responsibility requires that the defendant's behavior was performed with a requisite mental state. Some bodily movements are intentional and performed in a state of reasonably integrated consciousness. Some are not. Some defendants possess the requisite mental state, the intent to cause a prohibited harm such as death. Some do not. The truth of determinism does not entail that actions are indistinguishable from nonactions or that different mental states do not accompany action. These facts are true and make a perfectly rational legal difference even if determinism is true. Determinism is fully consistent with prima facie guilt and innocence.

Now consider the affirmative defenses of insanity and duress. Some people with a mental disorder do not know right from wrong. Others do. In cases of potential duress, some people face a hard choice that a person of reasonable firmness would yield to. These differences make perfect sense according to dominant retributive and consequential theories of punishment. A causal account can explain how these variations were caused, but it does not mean that these variations do not exist. Determinism is fully consistent with both the presence and absence of affirmative defenses. In sum, the legal criteria used to identify which defendants are criminally responsible map onto real behavioral differences that justify differential legal responses.

In their widely noted paper, Joshua Greene and Jonathan Cohen (2004) take issue with the foregoing account of the positive foundations of legal responsibility. They suggest that despite the law's official position, most people hold a dualistic, libertarian view of the necessary conditions for responsibility because "vivid scientific information about the causes of criminal behavior leads people to doubt certain individuals' capacity for moral and legal responsibility" (Greene and Cohen 2004, p. 1776). To prove their point, they use the hypothetical of "Mr. Puppet," a person who has been genetically and environmentally engineered to be a specific type of person. Greene and Cohen correctly point out that Mr. Puppet is really no different from an identical person I call Mr. Puppet2, who became the same sort of person without intentional intervention. Yet most people might believe that Mr. Puppet is not responsible. If so, however, should Mr. Puppet2 also not be responsible? After all, everyone is a product of a gene/environment interaction. But would it not then follow, as Greene and Cohen claim, that no one is responsible?

Greene and Cohen are correct about ordinary peoples' intuitions, but people make the fundamental psycholegal error (Morse 1994) all the time. That is, they hold the erroneous but persistent belief that causation is per se an excusing condition. This is a sociological observation and not a justification for thinking causation or determinism does or should excuse behavior. Whether the cause for behavior is biological, psychological, sociological, or astrological, or some frothy brew of all of these does not matter. In a causal universe, all behavior is presumably caused by its necessary and sufficient causes. A cause is just a cause. If causation excused behavior, no one could ever be responsible. Our law and morality do hold some people responsible and excuse others. Thus causation per se cannot be an excusing condition, no matter how much explanatory and predictive power a cause or set of causes for a particular behavior might have. The view that causation excuses per se is inconsistent with our positive doctrines and practices. Moreover, if Mr. Puppet and Mr. Puppet2 are both rational agents, the argument I have provided suggests that they are both justifiably held responsible. The lure of purely mechanistic thinking about behavior when causes are discovered is powerful but should be resisted.

At present, the law's "official" position about persons, action, and responsibility is justified unless and until neuroscience or any other discipline demonstrates convincingly that we are not the sorts of creatures we and the law think we are—conscious and intentional creatures who act for reasons that play a causal role in our behavior—and thus that the foundational facts for responsibility ascriptions are mistaken. If it is true, for example, that we are all automata, then no one is an agent, no one is acting and, therefore, no one can be responsible for action. But none of the stunning discoveries in the neurosciences or their determinist implications have yet begun to justify the belief that we are radically mistaken about ourselves. Let us therefore return to the proper understanding of the relation between neuroscience and law, again using criminal responsibility as the most powerful example.

The criteria for legal excuse and mitigation—like all legal criteria—are behavioral, including mental states. For example, lack of rational capacity is a generic excusing condition, which explains why young children and some people with mental disorder or dementia may be excused if they commit crimes. For another example, as Justice Oliver Wendell Holmes wrote long ago, "Even a dog distinguishes between being stumbled over and being kicked." Mental states matter to our responsibility for action. Take the insanity defense, for example, which excuses some people with mental disorder who commit crimes. The defendant will not be excused simply because he or she is suffering from mental disorder, no matter how severe it is. The defendant will not be excused simply because disordered thinking affected the defendant's reasons for action. Rather, the mental disorder must produce substantial lack of rational capacity concerning the criminal behavior in question. All insanity defense tests are primarily rationality tests. Lack of rational capacity is doing the excusing work.

Mental disorder that plays a role in explaining the defendant's behavior may paradoxically not have any effect on responsibility at all. Imagine a clinically hypomanic businessperson who, as a result of her clinical state, has really high attention, energy, and the like, and who makes a contract while in that state. If the deal turns out to be less advantageous than she thought, the law will not allow her to avoid that contract even though she made it under the influence of her mood disorder. Why? Because the businessperson was perfectly rational when she made the contract. Indeed, her hypomania might have made her "hyper-rational." Here is another example from criminal law. Imagine a person with paranoia who is constantly scanning his environment for signs of impending danger. Because the person is hypervigilant, he identifies a genuine and deadly threat to his life that ordinary people would not have perceived. If the person acts in self-defense, he is fully rational and his behavior would be justified. In this case, again, the mental abnormality made the agent "hyper-rational" in the circumstances.

Potentially legally relevant neuroimaging studies attempt to correlate brain activity with behavior, including mental states. In other words, legally relevant neuroscience must begin with behavior. We seek brain images associated with behaviors that we have already identified on normative, moral, political, and social grounds as important to us. For example, we recognize that adolescents behave differently from adults. They appear to be more impulsive and peer-oriented. They appear, on average, to be less fully rational than adults. These differences seemingly should make a moral and legal difference concerning, for example, criminal responsibility or the age at which people can drink or make independent health-care decisions. These differences also make us wonder if, in part, neuroanatomical or neurophysiological causal explanations might exist for the behavioral differences already identified as important to us.

Indeed, there is a parallel between the use of neuroscience for legal purposes and the development of cognitive neuroscience itself. Psychology does and must precede neuroscience when human behavior is in question (Hatfield, 2000).[2] Brain operations can be divided into various localities and subfunctions. The investigation of these constitutes the field of neuroscience. Some of the functions the brain implements are mental functions, such as perception, attention, memory, emotions, and planning. Psychology is broadly defined as the experimental science that directly studies mental functions. Therefore, psychology is the primary discipline investigating a major subset of brain functioning, including those functions that make us most distinctly human. These are also the types of functions that are therefore most relevant to law, because law is a human construction that is meant to help order human interaction. On occasion, inferring function from structure or physiology might be possible. In most cases, however, general knowledge or conjecture about function guides the investigation of structure and physiology. This will be especially true as we move "from the outside in." That is, it will be especially true as we study complex, intentional human behavior as opposed to, say, the perceptual apparatus. Lastly, therefore, psychology is the royal road to brain science in those areas that make us most distinctly human and that are most relevant to law.

When we evaluate what might be legally relevant brain science, we will be limited by the validity of the psychology upon which the brain science is based. As most—indeed, as all—honest neuroscientists and psychologists will admit, we wish that our psychological constructs and theories were better than they are. Thus, the legal helpfulness of neuroscience is limited.

Despite the limitations just described, neuroscience can sometimes be of assistance in helping us decide what a general rule should be and in adjudicating individual cases. Identifying brain correlates of legally relevant criteria is seldom necessary, or even helpful, when we are trying to define a legal standard if the behavioral difference is already clear. If the behavioral difference is not clear, then the neuroscience does not help, because the neuroscience must always begin with a behavior or behavioral difference that we have already identified as important.

For example, we have known that the rational capacity of adolescents is different from adults. Juvenile courts have existed for over a hundred years, well before anyone thought about neuroimaging the adolescent brain. The common law treated juveniles differently from adults for hundreds of years before we had any sense of neuroscience. People had to be of a certain age to vote, to drink, to join the army, and to be criminally responsible long before anyone envisioned functional magnetic resonance imaging (fMRI). If the rational capacity difference between adults and adolescents was less clear, then neuroscience could not tell us whether to treat adolescents differently, even if we believed that rationality made a difference. Whether adolescents are sufficiently different from adults so that they should be treated legally

2. What follows, in which I draw a parallel between the use of neuroscience for legal purposes and the development of cognitive neuroscience itself, borrows from and liberally paraphrases an excellent article by Hatfield (2000). What I will suggest is not meant to be critical or dismissive of neuroscience or of any other science. Indeed, I firmly believe that most neuroscience is genuinely excellent science. Nonetheless, much as legally relevant neuroscience must begin with identification of the behavior that is normatively relevant, so psychology does and conceptually must precede neuroscience.

differently is a behavioral and normative question in the first instance. Once the behavioral difference is established, at most the neuroscience concerning the biological immaturity of the adolescent prefrontal cortex does nothing more than provide a partial biological causal explanation of a normative relevant behavioral difference.

At this point one might object that the law draws bright, categorical lines when it is responding to behavioral continua and thus the law obscures important individual behavioral and moral differences. For example, although adolescents and adults *on average* demonstrate rationality differences that we think are morally and legally important, rationality is a continuum capacity and the adolescent and adult curves overlap, especially at the adolescent/adult margin. That is, some adults are less rational than some adolescents and some mid-to-late adolescents appear fully rational. Yet the law may create a bright-line category difference, as it does in the case of capital punishment. No capital killer who committed murder when he or she was sixteen or seventeen years old may be put to death, no matter how rational the adolescent may have been at the time. The law draws such bright lines because sometimes the costs of individualized decision making are too high and society is simply better off with a bright-line rule. This does not mean, however, that the law does not care about the behavioral differences when creating a general rule.

In what types of individual cases can neuroscience help? First, the data must be generally scientifically valid. They have to show precise correlations between brain states or activity and reliable and valid measures of legally relevant behavior, such as rational capacity. Such validity is increased if there is little overlap between the brain and behavior links of the groups being contrasted. In other words, the greater the overlap between the brain activity of people who do and do not meet a legal criterion, the greater will be the difficulty of using the scan of an individual to decide on which side of the line the person falls. And in nearly all cases, overlap will occur. Further, the technique and data must be valid for the individual case in question. For example, suppose the legal question is retrospective, such as determining a criminal defendant's mental state in the past at the time of the crime. Is a present scan a valid indication of what the defendant's mental state was during the criminal event?

Assume for the purposes of argument that we can solve both types of validity problems. If the behavioral evidence is clear, the neuroscience will be at most cumulative; it might have particular rhetorical force with a decision maker who is unsophisticated about how fMRI images are generated and the like. These images are not pictures of the brain, despite common belief that they are. Nonetheless, this evidence is superfluous. The expense of neuroimaging techniques and the inability successfully to use them with all potential subjects is reason not to use neuroscience in cases in which the behavior is clear. After all, the behavioral evidence is the most direct and probative evidence of behavioral criteria. Consequently, if we do use neuroimaging and the behavioral evidence and neuroscience evidence conflict, we must believe the behavioral evidence.

For example, if the neuroscientific evidence suggests that a criminal defendant has a rationality defect but the behavioral evidence indicates no defect whatsoever, we should conclude that the neuroscientific evidence is invalid in this case. Actions speak louder than images. Consider the following analogy. If a person does not complain of lower back pain or reduced mobility, we can safely conclude that she has no clinically significant lower back problem, even if an MRI of the spine shows substantial abnormalities. And, if the person shows clear signs of pain and reduced mobility, a problem exists even if the spine looks clean. Likewise, if the result of an IQ test does not accord with the subject's behavior, believe the behavior.

Here is an example from my own experience as a consultant forensic psychologist in a criminal case in which the defendant claimed that she was too unintelligent to be able to form the mental states required by the definition of the offense. She had taken apparently valid IQ tests indicating that her IQ was in the middle 60s—what is usually termed "mild retardation"—and no one could determine whether she was faking. My intervention was simple. I asked whether she had ever applied for a job for which she had to take some kind of screening test? Did she have kids in school, and if so, had she attended parent/teacher conferences? After all, teachers are really good at evaluating intelligence. The real-world data were collected and demonstrated without question that the defendant had an intelligence far above average.

To take a final example, suppose you were concerned with racial discrimination on the job. Research by Mahzarin Banaji, Elizabeth Phelps, and others has found, roughly speaking, that peoples' brain activity differentially responds to faces depending on whether the face presented is the same race or a different race from the subject. Should we assume a person is a racist or wrongly discriminates because his or her brain activates differentially? Suppose she has never expressed racist attitudes and her behavioral history shows that she always behaves equitably in the real world. Is the person a racist? Or, at least, for legal purposes, should we care about the differential brain activity?

What is the role of neuroevidence in cases in which the behavioral evidence is unclear or ambiguous? In such cases, valid neural markers and proxies may be extremely helpful evidence for determining whether a legal criterion is met. Neural indicators will rarely be dispositive because they will not be perfect markers, but they might genuinely be probative. On the other hand, the sensitivity or specificity of such markers might not be that high for specific legal criteria. If so, caution is warranted.

Now reconsider the case of the man who strangled his wife to death and threw her body out the window. He had a clear abnormality—a large benign subarachnoid cyst pressing on his frontal cortex—that was present at the time of the crime. Such a finding would certainly not be inconsistent with a rationality defect. But there was no hint that he suffered from a major mental disorder or had any substantial rationality defect either before or after the time of the crime. Some evidence existed that perhaps he had some impulse control problems, but nothing that had ever seriously interfered with his life or caused troubles with the law. Mild impulse control problems are not an excusing condition in any case. Sometimes otherwise law-abiding, rational, and responsible people just “lose” it. If the behavioral history had been more problematic, however, then the potential effect of the cyst might well have caused us more readily to believe that he did suffer from a rationality defect. Similarly, in a case in which intelligence is clearly relevant, if the real-world evidence about intelligence is ambiguous, scientific tests of intelligence, whether psychological or neuroscientific, would surely be helpful.

THE FUTURE OF IMAGING AND THE LAW

I have so far been arguing for a cautious, somewhat deflationary stance toward the potential of neuroscience to help us decide general and specific legal issues, but I do not mean to suggest I am a radical skeptic, cynic, or the like. I am not. I do not know what science is going to discover tomorrow. In the future we might well find grounds for greater optimism about broader legal relevance. But we have to be extraordinary sensitive to the limits of what neuroscience can contribute to the law at present.

How should the law proceed? What is the danger of images? The power of images to persuade might be greater than their legal validity warrants. First, images might not be legally relevant at all. We must always carefully ask what precise legal question is under consideration and then ask whether the image or any other neuroscience (or other type of) evidence actually helps us answer this precise question. The image might indicate something interesting, and it might be a vivid, compelling representation, but does it precisely answer our legal question?

Second, once naive subjects, such as average legislators, judges, and jurors, see images of the brain that appear correlated to the behavior in question, they tend to fall into the trap that I call the “lure of mechanism” or to make the fundamental psycholegal error discussed previously. That is, they tend to believe that causation is an excuse, especially if the brain seems to play a causal role. In a wonderful recent study (Knutson et al. 2007), researchers were able to predict with astonishing accuracy, depending on what part of the brain was active, whether the subject would or would not make a choice to buy a consumer item. The title of John Tierney’s article reporting on the study in the *New York Times*—an excellent example of an educated layperson’s response—asked, “The Voices in My Head Say ‘Buy It!’ Why Argue?” Tierney concluded, “You might remove the pleasure of shopping by somehow dulling the brain’s dopamine receptors . . . but try getting anyone to stay on that medication. Better the occasional jolt of pain. Charge it to the insula.” (Tierney 2007). Note the implication of mechanism: When you shop, you are not an acting agent but are at the mercy of your brain anatomy and physiology. You, the acting agent, the shopper, did not decide whether to buy. Your brain did. We are just brains in a mall. But we must resist the lure of mechanism. Brains do not shop; people do.

As a result, should we exclude imaging evidence from the courtroom, or should we, as we commonly do in the law, admit the evidence and trust cross-examination to expose its strengths and weaknesses? In other words, is the proper question the weight of such evidence or whether it should be admitted at all? The answer should depend on the relevance and strength of the science. If the legal relevance of the science is established and the science is quite good, my preference would be to admit the evidence and let the experts dispute its worth before the neutral adjudicator, the judge or the jury. But two criteria must be met first: The science must be both legally relevant and sound.

REFERENCES

Bok, H. 1998. *Freedom and Responsibility.* Princeton University Press.

Greene, J., and J. Cohen. 2004. For the law, Neuroscience changes nothing and everything. *Phil. Transactions of the Royal Society* B:1775–1785.

Hatfield, G. 2000. The brain's 'new' science: Psychology, neurophysiology, and constraint. *Philosophy of Science (Proceedings)* 67:S388-S403.

Kane, R. 2005. *A Contemporary Introduction to Free Will.* Oxford University Press.

Koenigs, M., L. Young, R. Adolphs, et al. 2007. Damage to the prefrontal cortex increases utilitarian judgments. *Nature* 446:908–911.

Knutson, B., S. Rick, G. E. Wimmer, D. Prelec, and G. Loewenstein. 2007. Neural predictors of purchases. *Neuron* 53:147–156.

Lenman, J. 2006. Compatibilism and contractualism: The possibility of moral responsibility. *Ethics* 117:7–31.

Mandavilli, A. 2006. Actions speak louder than images. *Nature* 466:664–665.

Morse, S. J. 2007. The non-problem of free will in forensic psychiatry and psychology. *Behavioral Sciences & the Law* 25:203–220.

Morse, S. J. 2004. Reason, results and criminal responsibility. *Illinois Law Review*:363–444.

Morse, S. J. 1994. Culpability and control. *University of Pennsylvania Law Review* 142:1587–1660.

People v. Weinstein, 154 Misc. 2d. 34 (NY 1992).

Roper v. Simmons, 543 U.S. 551 (2005).

Smilansky, S. 2000. *Free Will and Illusion.* Oxford University Press.

Strawson, P. F. 1982. Freedom and resentment. In *Free Will,* ed. G. Watson, 59–80. Oxford University Press.

Strawson, G. 1989. Consciousness, free will, and the unimportance of determinism. *Inquiry* 32:3–27.

Tierney, J. 2007, January 16. The voices in my head say 'Buy it!' Why argue? *New York Times* at F1.

Wallace, R. J. 1994. *Responsibility and the Moral Sentiments.* Harvard University Press.

CHAPTER 5

Neural Lie Detection in Courts

WALTER SINNOTT-ARMSTRONG

Getting scientists and lawyers to communicate with each other is not easy. Getting them to talk is easy, but communication requires mutual understanding—and that is a challenge.

LEGAL LINES ON SCIENTIFIC CONTINUA

Scientists and lawyers live in different cultures with different goals. Courts and lawyers aim at decisions, so they thrive on dichotomies. They need to determine whether defendants are guilty or not, liable or not, competent or not, adult or not, insane or not, and so on. Many legal standards implicitly recognize continua, such as when prediction standards for various forms of civil and criminal commitment speak of what is "highly likely" or "substantially likely," but in the end courts still need to decide whether the probability is or is not high enough for a certain kind of commitment. The legal system, thus, depends on on-off switches. This generalization holds for courts, and much of the legal world revolves around court decisions.

Nature does not work that way. Scientists discover continuous probabilities on multiple dimensions.[1] An oculist, for example, could find that a patient is able to discriminate some colors but not others to varying levels of accuracy in various circumstances. The same patient might be somewhat better than average at seeing objects far away in bright light but somewhat worse than average at detecting details nearby or in dim light. Given so many variations in vision, if a precise scientist were asked, "Is this particular person's vision good?" he or she could respond only with "It's good to this extent in these ways."

The legal system then needs to determine whether this patient's vision is good enough for a license to drive. That is a policy question. To answer it, lawmakers need to determine whether society can live with the number of accidents that are likely to occur if people with that level of vision get driver's licenses. The answer can be different for licenses to drive a car or a school bus or to pilot a plane, but in all such cases the law needs to draw lines on the continua that scientists discover.

1. This point is generalized from Fingarette (1972, 38–39).

The story remains the same for mental illness. Modern psychiatrists find large clusters of symptoms that vary continuously along four main dimensions.[2] Individual patients are more or less likely to engage in various kinds of behaviors within varying times in varying circumstances. For therapeutic purposes, psychiatrists need to locate each client on the distinct dimensions, but they do not need to label any client simply as insane or not.

Psychiatrists also need not use the terms "sane" and "insane" when they testify in trials involving an insanity defense. One example among many is the Model Penal Code test, which holds that a defendant can be found not guilty by reason of insanity if he lacks substantial capacity to appreciate the wrongfulness of his conduct or to conform his conduct to the requirements of the law. This test cannot be applied with scientific techniques alone. If a defendant gives correct answers on a questionnaire about what is morally right and wrong but shows no skin conductance response or activity in the limbic system while giving these answers, does that individual really "appreciate" wrongfulness? And does this defendant have a "capacity" to appreciate wrongfulness if he does appreciate it in some circumstances but not others? And when is that capacity "substantial"? Questions like these drive scientists crazy.

These questions mark the spot where science ends and policy decisions begin. Lawyers and judges can recognize the scientific dimensions and continua, but they still need to draw lines in order to serve their own purposes in reaching decisions. How do they draw a line? They pick a vague area and a terminology that can be located well enough in practice and that captures enough of the right cases for society to tolerate the consequences. Where lawmakers draw the line depends both on their predictions and on their values.

Courts have long recognized that the resulting legal questions can be confusing to psychiatrists and other scientists because their training lies elsewhere. Scientists have no special expertise on legal or policy issues. That is why courts in the past usually did not allow psychiatrists to testify on ultimate legal issues in trials following pleas of not guilty by reason of insanity. This restriction recently was removed in federal courts, but there is wisdom in the old ways, when scientists gave their diagnoses in their own scientific terms and left legal decisions to legal experts. In that system, scientists determine which dimensions are predictive and where a particular defendant lies on those continua. Lawyers then argue about whether that point is above or below the legal cutoff that was determined by judges or legislators using policy considerations. That system works fine as long as the players stick to their assigned roles.

This general picture applies not just to optometry and psychiatry but to other interactions between science and law, including neural lie detection. Brain scientists can develop neural methods of lie detection and then test their error rates. Scientists can also determine how much these error rates vary with circumstances, because some methods are bound to work much better in the lab than during a real trial. However, these scientists have no special expertise on the question of whether those error rates are too high to

2. These dimensions are standardized in American Psychiatric Association (2000).

serve as legal evidence. That is a policy question that depends on values; it is not a neutral scientific issue. This is one reason why neuroscientists should not be allowed to testify on the ultimate question of whether a witness is or is not lying.

Lying might appear different from insanity because insanity is a normative notion, whereas lying is not normative at all. A lie is an intentional deception without consent in order to induce reliance. Does the person who lies really believe that what he or she said is false? Well, he or she ascribes a probability that varies on a continuum. Does the speaker intend to induce belief and reliance? Well, that will not be clear if the plans are incomplete, indeterminate, or multiple. Does mutual consent exist, as in a game or some businesses? Well, varying degrees of awareness exist. Some cases are clear—maybe most cases. Nonetheless, what counts as a lie *is* partly a normative question that lies outside the expertise of scientists qua scientists. That is one reason why scientists should not be allowed to testify on that ultimate issue of lying. Their testimony should be restricted to their expertise.

FALSE NEGATIVES VERSUS FALSE POSITIVES

Although scientists can determine error rates for methods of lie detection, the issue is not so simple. For a given method in given circumstances, scientists distinguish two kinds of errors. The first kind of error is a false positive (or false alarm), which occurs when the test says that a person is lying but he or she really is not lying. The second kind of error is a false negative (or a miss), which occurs when the test says that a person is not lying but he or she really is lying. The rate of false positives determines the test's specificity, whereas the rate of false negatives determines the test's sensitivity.

These two error rates can differ widely. For example, elsewhere in this volume Nancy Kanwisher cites a study of one method of neural lie detection where one of the error rates was 31 percent and the other was only 16 percent. The error rate was almost twice as high in one direction than in the other. When error rates differ by so much, lawmakers need to consider each rate separately. Different kinds of errors create different problems in different circumstances. Lawmakers need to decide which error rate is the one that matters for each particular use of neural lie detection.

Compare three legal contexts: In the first a prosecutor asks the judge to let him use neural lie-detection techniques on a defense witness who has provided a crucial alibi for the defendant. The prosecutor thinks that this defense witness is lying. Here the rate of false positives matters much more than the rate of false negatives, because a false positive might send an innocent person to prison, and courts are and should be more worried about convicting the innocent than about failing to convict the guilty.

In contrast, suppose a defendant knows that he is innocent, but the trial is going against him, largely because one witness claims to have seen the defendant running away from the scene of the crime. The defendant knows that this witness is lying, so his lawyer asks the judge to let him use neural lie detection techniques on the accusing witness. Here the rate of false negatives matters more than the rate of false positives because a false negative is what might send an innocent defendant to prison.

Third, imagine that the defense asks the judge to allow as evidence the results of neural lie detection *on the accused* when he says that he did not commit the crime. Here the rate of false positives is irrelevant because the defendant would not submit this evidence if the results were positive for lying.

Overall, then, should courts allow neural lie detection? If the rates of false positives and false negatives turn out to differ widely (as I suspect they will), then the values of the system might best be served by allowing some uses in some contexts but forbidding others uses in other contexts. The legal system might not allow prosecutors to force any witness to undergo lie detection, but it still might allow prosecutors to use lie detection on some willing witnesses. Or the law might not allow prosecutors to use lie detection at all, but it still might allow defense attorneys to use lie detection on any witness or only on willing or friendly witnesses. If not even those uses are allowed, then the rules of evidence deprive the defense of a tool that, while flawed, could create a reasonable doubt, which is all the defense needs. If the intent is to ensure that innocent people are not convicted and if the defense volunteers to take the chance, then why the law should categorically prohibit this imperfect tool is unclear.

That judges would endorse such a bifurcated system of evidence is doubtful, although why is not clear. Some such system might turn out to be optimal if great differences exist between the rates of false negatives and false positives and also between the disvalues of convicting the innocent and failing to convict the guilty. Doctors often distinguish false positives from false negatives and use tests in some cases but not others, so why should courts not do the same? At least this question is worth thinking about.

BASE RATES

A more general problem, however, suggests that courts should not allow any neural lie detection. When scientists know the rates of false positives and false negatives for a test, they usually apply Bayes's theorem to calculate the test's positive predictive value, which is the probability that a person is lying, given a positive test result. This calculation cannot be performed without using a base rate (or prior probability). The base rate has a tremendous effect on the result. If the base rate is low, then the predictive value is going to be low as well, even if the rates of false negatives and of false positives seem reasonable.

This need for a base rate makes such Bayesian calculations especially problematic in legal uses of lie detection (neural or not). In lab studies the nature of the task or the instructions to subjects usually determines the base rate.[3] However, determining the base rate of lying in legal contexts is much more difficult.

3. For more on this, see Nancy Kanwisher's paper elsewhere in this volume.

Imagine that for a certain trial everyone in society were asked, “Did you commit this crime?” Those who answered “Yes” would be confessing, so almost everyone, including the defendant, would answer “No.” Only the person who was guilty would be lying. Thus, the base rate of lying in the general population for this particular question is extremely low. Hence, given Bayes’s theorem, the test of lying might seem to have a low predictive value.

However, this is not the right way to calculate the probability. What really needs to be known is the probability that someone is lying, given that this person is a defendant in a trial. How can that base rate be determined? One way is to gather conviction rates and conclude that most defendants are guilty, so most of them are lying when they deny their guilt. With this assumption, the base rate of lying is high, so Bayes’s theorem yields a high predictive value for a method of lie detection with low enough rates of false negatives and false positives. However, this assumption that most defendants are guilty violates important legal norms. Our laws require us to presume that each defendant is innocent until proven guilty. Thus, if a defendant is asked whether he did it and he answers, “No,” then our judicial system is legally required to presume that he is not lying. The system should not, then, depend on any calculation that assumes guilt or even a high probability of guilt. But without some such assumption, one cannot justify a high enough base rate to calculate a high predictive value for any method of neural lie detection of defendants who deny their guilt.

CONCLUSION

A crystal ball would be needed to conclude that neural lie detection has no chance of ever working or of being fair in trials. But many details need to be carefully worked out before such techniques should be allowed in courts. Whether the crucial issues can be resolved remains to be seen, but the way to resolve them is for scientists and lawyers to learn to work together and communicate with each other.

REFERENCES

American Psychiatric Association. 2000. *The diagnostic and statistical manual of the American Psychiatric Association*, fourth edition, revised. Arlington, VA: American Psychiatric Publishing.

Fingarette, H. 1972. *The meaning of criminal insanity.* Berkeley: University of California Press.

Lie Detection in the Courts: The Vain Search for the Magic Bullet

JED S. RAKOFF

The detection of truth is at the heart of the legal process.[1] The purpose of a trial, in particular, is to resolve the factual disputes on which a case turns, and the trial culminates with the rendering of a "verdict," which in Latin means "to state the truth."

Given the common tendency of witnesses to exaggerate, to embroider, and, frankly, to lie,[2] how does a fact finder determine the truth? Particularly in the adversarial system of justice common to the Anglo-American tradition, truth is revealed, and lying detected, by exposing witnesses to cross-examination; that is, to tough questioning designed to test the consistency and credibility of the witness's story. John Henry Wigmore, known to most lawyers as the most profound expositor of the rules of evidence in the history of law, famously described cross-examination as "the greatest legal engine ever invented for the discovery of truth."[3] And while not everyone is specially trained in the art of cross-examination, common experience—whether it be of parents questioning children, of customers questioning salespersons, or of reporters questioning politicians—suggests that nothing exposes fabrication like a good tough questioning.

But no one supposes that cross-examination is a perfect instrument for detecting the truth. For that matter, there probably has never been a scientific study of how effective cross-examination is in detecting lies, and I am

1. This article, based on remarks made at the American Academy of Arts and Sciences's conference on February 2, 2007, presents the author's personal views and does not reflect how he might decide any legal issue in any given case.

2. In my remarks on which this article is based, I estimated that 90 percent of all material trial witnesses knowingly lie in some respect (though not always materially). This estimate is based on my experience as a trial judge for the past twelve years and as a trial lawyer for the twenty-five years previous to that and has no scientific study to back it up. Possibly I am too harsh: perhaps the percentage of material witnesses who consciously lie in some respect is as low as, say, 89 percent.

3. John Henry Wigmore, Evidence In Trials At Common Law (Chadbourn Edition, 1974), Vol. 5, p. 32.

not even sure how such a study could be devised. In any event, people have always sought for some simpler, talismanic way of separating truth from falsehood, and, relatedly, of determining guilt or innocence.

Thus, in medieval times, an accused who protested his innocence was put to such tests as the ordeal by water, in which he was bound and thrown in the river. The accused who had falsely protested his innocence would be rejected by the water and would float, whereas the honest accused would be received by the water and immediately sink. Usually he was fished out before he drowned; but, if the rescue came too late, he at least died in a state of innocence. If one accepted the basic religious theories on which these tests were premised, the tests were infallible.

As the middle ages gave way to the Renaissance, the faith-based methods of the ordeals were replaced by a more up-to-date, empirical method for determining the truth: torture. Although some men and women, in their perversity, might deny the evil of which their accusers were certain they were guilty, infliction of sufficient pain would lead them to admit, mirabile dictu, exactly what the accusers had suspected.

After a while, however, it became increasingly obvious that torture was leading to too many "false positives," and more accurate methods were sought. In his treatise on evidence, Wigmore contends that cross-examination was originally developed as an alternative to torture.[4] Today, of course, it is inconceivable that anyone would recommend the use of torture.

From the seventeenth century onward, cross-examination, for all its limitations, seemed to be the best the legal system had to offer as a means of determining the truth. In the late nineteenth century, every schoolchild knew the story of how Abe Lincoln, in defending a man accused of murder, carefully questioned the prosecution's eyewitness as to how he was able to see the accused commit the murder in the dead of night and, when the witness said it was because there was a full moon, produced an almanac showing that on that night the moon was but a sliver. As Lincoln elsewhere said, "you can't fool all of the people all of the time"—at least not when someone is around to ask the hard questions.

But the late nineteenth century also witnessed the growing belief that all areas of inquiry would ultimately yield to the power of science. It was just a matter of time before an allegedly "scientific" instrument for detecting lies was invented, namely, the polygraph.

The truth is that the polygraph is not remotely scientific. The theory of the polygraph—itself largely untested—is that someone who is consciously lying feels anxiety, and that that anxiety, in turn, is manifested by an increase in respiration rate, pulse rate, blood pressure, and sweating. Common experience suggests many possible flaws in this theory: more practiced liars might feel little anxiety about lying; taking a lie detector test might itself generate anxiety; sweating, pulse rate, blood pressure, and respiration rate are commonly affected by all sorts of other conditions, both external and internal; and so forth. One might hypothesize, therefore, that polygraph tests, while they might be better than pure chance in separating truth tellers from liars—

4. Ibid., p. 32ff.

after all, some people might fit the theory—would nevertheless have a high rate of error. As Marcus E. Raichle discusses elsewhere in this volume, that is precisely what the National Academies (NAS), which in 2002 reviewed the evidence on polygraph reliability, concluded. The NAS also concluded that polygraph testing has “weak scientific underpinnings”[5] and that “belief in its accuracy goes beyond what the evidence suggests.”[6]

Not all experts agree. Reviewing the literature in 1998, the Supreme Court of the United States concluded that “the scientific community remains extremely polarized about the reliability of polygraph techniques,” with some studies concluding that polygraphs are no better than chance at detecting lies and, at the other extreme, one study concluding that polygraph results are accurate about 87 percent of the time. But even a 13 percent error rate is a high number when you are dealing with something as important as determining a witness’s credibility, let alone determining whether he or she is guilty or innocent of a crime.

Moreover, all these error-rate statistics are suspect because the scientific community is nowhere close to agreeing on how one properly establishes the base measure for determining the reliability of the polygraph. To devise an experiment in which one set of subjects is told to lie and the other set of subjects is told to tell the truth is one thing; to recreate the real-life conditions that would allow for a true test of the polygraph is quite something else. Whether any sound basis exists on which one can assert anything useful about the reliability or unreliability of the polygraph is uncertain.

Courts, being conservative and skeptical by nature, have largely tended to exclude polygraph evidence. But that has not stopped the government, the military, some private industry, and much of the public generally from accepting the polygraph as reliable—so great is the desire for a “magic bullet” that can instantly distinguish truth from falsehood.

Even the courts, while excluding polygraph evidence from the courtroom, have sometimes approved its use by the police on the cynical basis that it really does not matter whether the polygraph actually detects lying, so long as people believe that it does: if a subject believes that a polygraph actually works, he or she will be motivated to tell the truth and “confess.” The hypocrisy of this argument is staggering: the argument, in effect, is that even if the truth is that polygraph tests are, at best, error-prone, the police and other authorities should lie to people and encourage them to believe that the tests are highly accurate because this lie will encourage people to tell the truth.

Even on these terms, moreover, experience in my own courtroom suggests that the use of polygraphs is much more likely to cause mischief, or worse, than to be beneficial. Let me give just one example. The Millenium Hotel is situated next to Ground Zero. A few weeks after the attack on the Twin Towers, hotel employees were allowed back into the hotel to recover the belongings of the guests who had had to flee the premises on September 11, and one

5. The National Academies, National Research Council, “Polygraph Testing Too Flawed for Security Screening,” October 8, 2002, p. 2.

6. The National Academies, National Research Council, Committee to Review the Scientific Evidence on the Polygraph, The Polygraph and Lie Detection (2003), p. 7.

of the hotel's security guards reported to the Federal Bureau of Investigation (FBI) that he had found in the room safe on the fiftieth floor, in a room occupied by a man named Abdullah Higazy, a copy of the Koran and a pilot's radio of the kind used to guide planes from the ground. The FBI quickly discovered that Higazy was a former member of the Egyptian Air Force now resident in Brooklyn, but when they questioned him, he denied ever having a pilot's radio. Hypothesizing that he was lying to cover up his use of the radio to guide the terrorist pilots to the Twin Towers, the FBI arrested Higazy and brought him before me on a material witness warrant. At the hearing, Higazy repeatedly asked to be given a polygraph test to establish that the radio was not his. I explained to him that polygraph tests were too unreliable to be admitted in court. Nevertheless, after the hearing, Higazy, over his own lawyer's recommendation, asked the FBI to give him a polygraph test.

The FBI brought Higazy, alone, into the polygraph testing room, explaining that his lawyer could not be present because it would upset the balance of this "delicate" test. Over the next three hours, the FBI agent administering the test repeatedly told Higazy that he was not being truthful. Finally, Higazy, by now hysterical, blurted out that maybe the radio really was his. At that point, the FBI stopped the test and told the lawyer that Higazy had confessed and would be charged, at a minimum, with making false statements to the FBI and possibly with aiding and abetting the attack on the Twin Towers, a capital offense. The next day, based on the prosecutor's flat statement that Higazy had confessed, I ordered Higazy detained without bail and he was shortly thereafter formally charged with lying to the FBI.

Three days later, an American Airlines pilot contacted the Millennium Hotel and asked if he could get back the pilot's radio he had left there on September 11. It quickly developed that the radio was, indeed, his and had never been in Higazy's room or possession. The Millennium security guard had made up the whole story about finding the radio in Higazy's room, apparently because he wanted revenge for 9/11 on anyone who had Arab ancestry. The government dropped the charges against Higazy and prosecuted the security guard instead, who pled guilty to lying to the FBI.

For my part, I ordered an investigation by the government into the circumstances of the FBI's polygraph testing, the result of which was a report assuring me that the manner and mode of Higazy's polygraph examination was consistent with standard FBI practice. I am not sure whether this means that the FBI really believes in its polygraph results, despite their inaccuracy, or whether the FBI simply uses the façade of polygraph testing to try to elicit confessions. Either way, but for a near miracle, Mr. Higazy might likely now either be rotting in prison or facing execution.

Why have I spent so much time describing the evils of polygraphs, when the primary topic of this volume is the brave new world of brain scanning? I believe that many of the same evils are likely to result from the use of brain scanning to detect lies unless we are very, very careful. If anything, the potential for mischief is even greater, because while polygraphy was largely developed by technicians, brain scanning as a means of detecting lies is said to be the product of studies by honest-to-goodness real-life neuroscientists.

But the credentials of the scientists should not obscure the shakiness of the science.[7] A basic problem with both polygraphy and brain scanning to detect lying is that no established standard exists for defining what willful deception is, let alone how to establish a base measure against which the validity and reliability of any lie-detection technique can be evaluated. What exists at this point are imaging technologies that show us patterns of activity or other events in the brain that are hypothesized to correlate with various mental states. Not one of these hypotheses has been subjected to the kind of rigorous testing that would establish its validity.

That, however, has not stopped several commercial enterprises from offering brain scanning as a purportedly scientific lie-detection technique that law enforcement agencies, private businesses, and even courts should utilize. The mere fact that evidence is proffered by someone with scientific credentials does not begin to satisfy the conditions for its admissibility in court.

In the case of the federal courts, the admissibility of expert testimony is governed by Rule 702 of the Federal Rules of Evidence. Rule 702 provides that

> If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training or education may testify thereto in the form of an opinion of otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

Though every case must be assessed on its individual merits, brain scanning as a means of assessing credibility likely suffers from several defects that would render such evidence inadmissible under Rule 702 as it has been interpreted in the federal courts.

First, and perhaps most fundamentally, there is no commonly accepted theory of how brain patterns evidence lying: in the absence of such a theory, all that is being shown, at best, is the presence or absence of certain brain patterns that allegedly correlate with some hypothesized accompaniment of lying, such as anxiety. But no scientific evidence has shown either that lying is always accompanied by anxiety or that anxiety cannot be caused by a dozen other factors that cannot be factored out.

Second, the theories that have been proposed have not been put to the test of falsifiability, which, if one accepts (as the Supreme Court does) a Popper-like view of science, is the sine qua non of assessing scientific validity and reliability.

Third, no standard way exists of defining what lying is, let alone how to test for it. The law recognizes many kinds of lies, ranging from "white lies" and "puffing" to affirmative misstatements, actionable half-truths, and material omissions. Brain scans cannot yet come close to distinguishing between these different kinds of lying. Yet the differences are crucial in almost any case: a little white lie is altogether different, in the eyes of the law and of common

7. For a more expert discussion of the limitations of brain scanning as a truth-detection device, see the articles by Elizabeth Phelps and Nancy Kanwisher elsewhere in this volume.

sense, from an intentional scheme to defraud. Nothing in the brain-scan approach to lie detection even attempts to make such distinctions. And what might a brain scan be predicted to show in the case of a lie by omission; that is, the person whose statements are truthful as far as they go but who conceals a material fact that puts an entirely different perspective on what is being said? In my experience, these are the most common kinds of lies in court, and they are revealed only by a good cross-examination.

For these and other reasons, brain-scanning tests of credibility might well fail to meet the tests for admissibility under Rule 702, both because the brain-scanning tests lack scientific reliability and because they are unlikely to be useful to the jury. Additionally, even evidence that qualifies for admission under Rule 702 may be excluded under Rule 403 of the Federal Rules of Evidence, which provides that "Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undo delay, waste of time, or needless presentation of cumulative evidence." Brain-scanning evidence, precisely because it holds itself out as truly scientific, is likely to have a much greater impact on the trier of fact than its limited theoretical and experimental bases can fairly support; it therefore might also be excluded under Rule 403.

A time may come when some sort of brain-scanning technique will be developed that will actually show the act of willful lying with enough accuracy and predictability as to meet both the requirements of law and the even more rigorous standards of accepted science. But long before that occurs, claims will be made—some even by reputable scientists whose enthusiasm for their own studies has overwhelmed all caution—that science has developed brain-scan lie detectors that are perfectly accurate, when in fact they are not. Indeed, some such claims are already being made. The desire for a magic bullet that exposes lies is so strong that many people will be persuaded that, indeed, brain scans can achieve what polygraphers have long claimed but, in my view, wholly failed to achieve. The result will be an abdication of the difficult work of actually detecting lies through cross-examination in favor of quasi-scientific tests that substitute the façade of scientific certainty for the actuality of truth.

Neuroscience-Based Lie Detection: The Need for Regulation

HENRY T. GREELY

"I swear I didn't do it."
"The check is in the mail."
"The article is *almost* done."

In our lives and in our legal system we often are vitally interested in whether someone is telling us the truth. Over the years, humans have used reputation, body language, oaths, and even torture as lie detectors. In the twentieth century, polygraphs and truth serum made bids for widespread use. The twenty-first century is confronting yet another kind of lie detection, one based on neuroscience and particularly on functional magnetic resonance imaging (fMRI).

The possibility of effective lie detection raises a host of legal and ethical questions. Evidentiary rules on scientific evidence, on probative value compared with prejudicial effect, and, possibly, rules on character evidence would be brought into play. Constitutional issues would be raised under at least the Fourth, Fifth, and Sixth Amendments, as well as, perhaps, under a First Amendment claim about a protected freedom of thought. Four U.S. Supreme Court justices have already stated their view that even a perfectly effective lie detector should not be admissible in court because it would unduly infringe the province of the jury. And ethicist Paul Wolpe has argued that this kind of intervention raises an entirely novel, and deeply unsettling, ethical issue about privacy within one's own skull.[1]

These issues are fascinating and the temptation is strong to pursue them, but we must not forget a crucial first question: does neuroscience-based lie detection work and, if so, how well? This question has taken on particular urgency as two, and possibly three, companies are already marketing fMRI-based lie detection services in the United States. The deeper implications of effec-

1. Paul R. Wolpe, Kenneth R.. Foster, and David D. Langleben, "Emerging Neurotechnologies for Lie-Detection: Promises and Perils," *American Journal of Bioethics* 5 (2) (2005): 39–49.

tive lie detection are important but may prove a dangerous distraction from the preliminary question of effectiveness. Their exploration may even lead some readers to infer that neuroscience-based lie detection is ready for use.

It is not. And nonresearch use of neuroscience-based lie detection should not be allowed until it has been proven safe and effective.[2] This essay will review briefly the state of the science concerning fMRI-based lie detection; it will then describe the limited extent of regulation of this technology and will end by arguing for a premarket approval system for regulating neuroscience-based lie detection, similar to that used by the Food and Drug Administration to regulate new drugs.

NEUROSCIENCE-BASED LIE DETECTION: THE SCIENCE

Arguably all lie detection, like all human cognitive behavior, has its roots in neuroscience, but the term *neuroscience-based lie detection* describes newer methods of lie detection that try to detect deception based on information about activity in a subject's brain.

The most common and commonly used lie detector, the polygraph, does not measure directly activity in the subject's brain. From its invention around 1920, the polygraph has measured physiological indications that are associated with the mental state of anxiety: blood pressure, heart rate, breathing rate, and galvanic skin response (sweating). When a subject shows higher levels of these indicators, the polygraph examiner may infer that the subject is anxious and further that the subject is lying. Typically, the examiner asks a subject a series of yes or no questions while his physiological responses are being monitored by the device. The questions may include irrelevant questions and emotionally charged, "probable lie" control questions as well as relevant questions. An irrelevant question might be "Is today Tuesday?" A probable lie question would be "Have you ever stolen anything?" a question the subject might well be tempted to answer "no," even though it is thought unlikely anyone could truthfully deny *ever* having stolen anything. Another approach, the so-called guilty knowledge test, asks the subject questions about, for example, a crime scene the subject denies having seen. A subject who shows a stronger physiological reaction to a correct statement about the crime scene than to an incorrect statement may be viewed as lying about his or her lack of knowledge.

The result of a polygraph examination combines the physiological results gathered by the machine with the examiner's assessment of the subject to draw a conclusion about whether the subject answered particular questions honestly. The problems lie in the strength (or weakness) of the connection between the physiological responses and anxiety, on the one hand, and both anxiety and deception, on the other. Only if both connections are powerful can one argue that the physiological reactions are strong evidence of deception.

2. This argument is made at great length in Henry T. Greely and Judy Illes, "Neuroscience-Based Lie Detection: The Urgent Need for Regulation," *American Journal of Law & Medicine* 33 (2007): 377–431

A 2003 National Academy of Sciences (NAS) report analyzed polygraphy in detail. The NAS found little rigorous scientific assessment of polygraph accuracy but concluded that in good settings it was substantially better than chance—and substantially less than perfect. The NAS further noted that subjects could take plausible countermeasures to lower the device's accuracy even further. The NAS advised against the use of polygraphs for personnel screening.

Researchers are now working on at least five methods to produce a newer generation of lie detection devices, devices that measure aspects of the brain itself rather than the physiological responses associated with anxiety. One approach uses electroencephalography to look for a so-called P300 wave in the brain's electrical activity. This signal, seen about 300 milliseconds after a stimulus, is said to be a sign that the person (or the person's brain) recognizes the stimulus. A second method uses near-infrared laser spectroscopy to scatter a laser beam off the outer layers of the subject's brain and then to correlate the resulting patterns with deception. Proponents of a third method, periorbital thermography, claim to be able to detect deception by measuring an increase in the subject's temperature around the eyes, allegedly as a result of increased blood flow to the prefrontal regions. A fourth approach analyzes fleeting "facial micro-expressions" and is more like the polygraph in that it seeks assertedly involuntary and uncontrollable body reactions correlated with deception rather than looking at direct measures of aspects of the brain. That these new methods will work is far from clear. Almost no peer-reviewed literature exists for any of them.

The most advanced neuroscience-based method for lie detection is fMRI. As described in more detail by Marcus Raichle elsewhere in this volume, fMRI uses magnetism to measure the ratio of oxygenated to deoxygenated hemoglobin in particular areas of the brain, three-dimensional regions referred to as "voxels." The Blood Oxygenation Level Dependence (BOLD) hypothesis holds that a higher ratio of oxygenated to deoxygenated blood in a particular voxel correlates to higher energy consumption in that region a few seconds earlier. An fMRI scan can document how these ratios change as subjects perceive, act, or think different things while being scanned. Then sophisticated statistical packages look at the changes in thousands of voxels to find correlations between these blood oxygen ratios and what was happening to the subject several seconds earlier.

Since the development of fMRI in the early 1990s, many thousands of peer-reviewed papers have been published using the technique to associate particular patterns of blood flow (and, hence, under the BOLD hypothesis, brain activity) with different mental activities. Some of these associations have been entirely plausible and have been adopted in neurosurgery; for example, using fMRI to locate precisely the location of a particular region of a patient's brain in order to guide the surgeon's scalpel. Other claimed connections are more surprising, like using fMRI to locate brain regions responsible for passionate, romantic love or for a nun's feeling of mystical union with God. Still other published associations are about lie detection.

As of March 2007, at least twelve peer-reviewed articles from eight different laboratories had been published on fMRI-based lie detection.[3] The different laboratories used different experimental designs (sometimes the same laboratory used different designs in different publications), but each claimed to find statistically significant correlations between deception and certain patterns of brain activity. This apparent scientific support for fMRI-based lie detection becomes much weaker on examination. This body of work has at least six major flaws.

First, almost none of the work has been replicated. One of the laboratories, Andrew Kozel's, replicated at least one of its own studies and two of Daniel Langleben's published studies are quite similar, though not identical.[4] None of the other laboratories has replicated, at least in the published literature, their own studies. More important, none of the studies has been replicated by other labs. Replication is always important in science; it is particularly important with a new and complex technology like fMRI, where anything from the details of the experimental design, the method of subject selection, or the technical aspects of the individual MRI machine on any particular day can make a great difference.

Second, although the studies all find associations between deception and activation or deactivation in some brain regions, they often disagree among themselves in what brain regions are associated with deception. This some-

3. Sean A. Spence et al., "Behavioral and Functional Anatomical Correlates of Deception in Humans," *Brain Imaging Neuroreport* (2001): 2849; Tatia M. C. Lee et al., "Lie Detection by Functional Magnetic Resonance Imaging," *Human Brain Mapping* 15 (2002): 157 (manuscript was received by the journal two months before Spence's earlier-published article had been received and, in that sense, may be the earliest of these experiments); Daniel D. Langleben et al., "Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study," *Neuroimage* 15 (2002): 727; G. Ganis et al., "Neural Correlates of Different Types of Deception: An fMRI Investigation," *Cerebral Cortex* 13 (2003): 830; F. Andrew Kozel et al., "A Pilot Study of Functional Magnetic Resonance Imaging Brain Correlates of Deception in Healthy Young Men," *Journal of Neuropsychiatry & Clinical Neuroscience* 16 (2004): 295; F. Andrew Kozel, Tamara M. Padgett, and Mark S. George, "Brief Communications: A Replication Study of the Neural Correlates of Deception," *Behavioral Neuroscience* 118 (2004): 852; F. Andrew Kozel et al., "Detecting Deception Using Functional Magnetic Imaging," *Biological Psychiatry* 58 (2005): 605; Daniel D. Langleben et al., "Telling Truth from Lie in Individual Subjects with Fast Event-Related fMRI," *Human Brain Mapping* 26 (2005): 262; C. Davatzikos et el., "Classifying Spatial Patterns of Brain Activity with Machine Learning Methods: Application to Lie Detection," *Neuroimage* 28 (2005): 663; Tatia M. C. Lee et al., "Neural Correlates of Feigned Memory Impairment," *Neuroimage* 28 (2005): 305; Jennifer Maria Nunez et al., "Intentional False Responding Shares Neural Substrates with Response Conflict and Cognitive Control," *Neuroimage* 25 (2005): 267; Feroze B. Mohamed et al., "Brain Mapping of Deception and Truth Telling about an Ecologically Valid Situation: Function MR Imaging and Polygraph Investigation–Initial Experience," *Radiology* 238 (2006): 679.

4. Kozel, Padgett, and George, "Brief Communications: A Replication Study of the Neural Correlates of Deception," replicates Kozel et al., "A Pilot Study of Functional Magnetic Resonance Imaging Brain Correlates of Deception in Healthy Young Men." Langleben's first study, Langleben et al., "Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study," is closely mirrored by his second and third, Langleben et al., "Telling Truth from Lie in Individual Subjects with Fast Event-Related fMRI," and Davatzikos et al., "Classifying Spatial Patterns of Brain Activity with Machine Learning Methods: Application to Lie Detection."

times happens within one laboratory: Langleben's first two studies differed substantially in what regions correlated with deception.[5]

Third, only three of the twelve studies dealt with predicting deceptiveness by individuals.[6] The other studies concluded that on average particular regions in the (pooled) brains of the subjects were statistically significantly likely to be activated (high ratio of oxygenated to deoxygenated hemoglobin) or deactivated (low ratio) when the subjects were lying. These group averages tell you nothing useful about the individuals being tested. A group of National Football League place kickers and defensive linemen could, on average, weigh 200 pounds when no single individual was within 80 pounds of that amount. The lie-detection results are not likely to be that stark, but before we can assess whether the method might be useful, we have to know how accurate it is in detecting deception by individuals—its specificity (lack of false positives) and sensitivity (lack of false negatives) are crucial. Only one of the Kozel articles and two of the Langleben articles discuss the accuracy of individual results.

Next is the question of the individuals tested. The largest of these studies involved thirty-one subjects;[7] more of them looked at ten to fifteen. The two Langleben studies that looked at individual results were based on four subjects. For the most part, the subjects were disconcertingly homogenous—young, healthy, and almost all right-handed. Langleben's studies, in particular, were limited to young, healthy, right-handed undergraduates at the University of Pennsylvania who were not using drugs. How well these results project to the rest of the population is unknown.

A fifth major problem is the artificiality of the experimental designs. People are recruited and give their informed consent to participate in a study of fMRI and deception. Typically, they are told to lie about something. In Langleben's three studies they were told to lie when they saw a particular playing card projected on the screen inside the scanner. In Kozel's work, perhaps the least artificial of the experiments, subjects were told to take either a ring or a watch from a room and then to say, in the scanner, that they had not taken either object. Note how different this is from a criminal suspect telling the police that he had not taken part in a drug deal, or, for that matter, from a dinner guest praising an overcooked dish. The experimental subjects are following orders to lie where nothing more rides on the outcome than (in some cases) a promised $50 bonus if they successfully deceive the researchers. We just do not know how well these methods would work in settings similar to those where lie detection would, in practice, be used.

5. Compare Langleben et al., "Brain Activity During Simulated Deception: An Event-Related Functional Magnetic Resonance Study," with Langleben et al., "Telling Truth from Lie in Individual Subjects with Fast Event-Related fMRI."

6. Langleben et al., "Telling Truth from Lie in Individual Subjects with Fast Event-Related fMRI"; Davatzikos et al., "Classifying Spatial Patterns of Brain Activity with Machine Learning Methods: Application to Lie Detection"; and Kozel et al., "Detecting Deception Using Functional Magnetic Imaging."

7. Kozel et al., "Detecting Deception Using Functional Magnetic Imaging."

Finally, and perhaps most worryingly, as with the polygraph, countermeasures could make fMRI-based lie detection ineffective against trained liars. And countermeasures are easy with fMRI. One can ruin a scan by movement of the head or sometimes of just the tongue. Or, more subtly, as the scanner is detecting patterns of blood flow associated with brain activity, one can add additional brain activity. What happens to these results if the subject, when answering, is also reciting to himself the multiplication tables? We have no idea.

The few published papers that have looked at individuals have claimed accuracy rates of about 70 to around 90 percent in detecting lies. These results —not substantially different, by the way, from reported results with the polygraph—must be taken with a grain of salt. We just do not know how reliably accurate fMRI-based lie detection will be with diverse subjects in realistic settings, with or without countermeasures. For now, at least, based on the peer-reviewed literature, the scientific verdict on fMRI-based lie detection seems clear: interesting but not proven.

NEUROSCIENCE-BASED LIE DETECTION: THE LAW

In spite of this lack of convincing proof of efficacy, at least two companies in the United States—No Lie MRI and Cephos Corp—are offering fMRI-based lie detection services. They can do this because of the near absence of regulation of lie detection in the United States.

In general, the use of lie detectors is legal in the United States. The polygraph is used thousands of times each week for security screenings, in criminal investigations, as part of the conditions of release for sex offenders. The device can be used even more broadly, subject to almost no regulation, with one major exception: employers.

The federal Employee Polygraph Protection Act (EPPA) of 1988 forbids most employers from forcing job applicants and most employees to take lie-detector tests and from using the results of such tests. As a result, no American can today legally face, as I did, a polygraph test when applying, at age twenty-one, for a job as a bartender at a pizza parlor. Some employers are granted exceptions, notably governments and some national security and criminal-investigation contractors. The act's definition of lie detection is broad (although No Lie MRI has frivolously argued that EPPA does not apply to fMRI). The act also exempts the use of polygraphs (not other forms of lie detection) on employees in some kinds of employer investigations, subject only to some broad rights for the employees. About half the states have passed their own versions of this act, applying it to most or all of their state and local employees. Some states have extended protections against lie detection to a few other situations, including in connection with insurance claims, welfare applications, or credit reports. (A few states have required the use of lie detection in some settings, such as investigations of police officers.)

Almost half of states have a licensing scheme for polygraph examiners. A few of these statutes may effectively prohibit fMRI-based lie detection because they prohibit lie detection except by licensed examiners and provide only for licensing polygraph examiners, not fMRI examiners. No state, however, has yet explicitly regulated neuroscience-based lie detection.

One site for the possible use of lie-detection technology is particularly sensitive—the courtroom. Thus far, fMRI-based lie detection has not been admitted into evidence in court. The courts will apply their own tests in making such decisions. However, the eighty-plus years of litigation over courtroom uses of polygraph evidence might provide some useful lessons.

The polygraph is never admissible in U.S. courtrooms—except when it is. Those exceptions are few but not trivial. In state courts in New Mexico, polygraph evidence is presumptively admissible. In every other American state and in the federal courts, polygraph evidence is generally not admissible. Some jurisdictions will allow it to be introduced to impeach a witness's credibility. Others will allow its use if both parties have agreed, before the test was taken, that it should be admitted. (This willingness to allow polygraph to be admitted by the parties' stipulation has always puzzled me; should judges allow the jury to hear, as scientific evidence, the results of palm reading or the Magic Eight Ball if the parties stipulated to it?) At least one federal court has ruled that a defendant undergoing a sentencing hearing where the death penalty may be imposed is entitled to use polygraph evidence to try to mitigate his sentence.[8]

U.S. courts have rejected the polygraph on the grounds that it is not acceptable scientific evidence. For many years federal and state courts used as the test of admissibility of scientific evidence a standard taken from the 1923 case *Frye v. United States* (293 F. 1013 [DC Cir 1923]), which involved one of the precursors to the polygraph. *Frye,* which required proof that the method was generally accepted in the scientific community, was replaced in the federal courts (and in many state courts) by a similar but more complicated test taken from the 1993 U.S. Supreme Court case, *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (509 U.S. 579 [1993]). Both cases fundamentally require a finding that the evidence is scientifically sound, usually based on testimony from experts. Under both *Frye* and *Daubert,* American courts (except in New Mexico) have uniformly found that the polygraph has not been proven sufficiently reliable to be admitted into evidence. Evidence from fMRI-based lie detection will face the same hurdles. The party seeking to introduce it at a trial will have to convince the judge that it meets the *Frye* or *Daubert* standard.

The U.S. Supreme Court confronted an interesting variation on this question in 1998 in a case called *United States v. Scheffer* (523 U.S. 303 [1998]). Airman Scheffer took part in an undercover drug investigation on an Air Force base. As part of the investigation, the military police gave him regular polygraph and urine tests to make sure he was not misusing the drugs himself. He consistently passed the polygraph tests but eventually failed the urine test, leading to his charge and conviction by court-martial.

8. *Rupe v. Wood*, 93 F.3d 1434 (9th Cir. 1996). *See also Height v. State*, 604 S.E.2d 796 (Ga. 2004).

Unlike the general Federal Rules of Evidence, the Military Rules of Evidence expressly forbid the admission of any polygraph evidence. Airman Scheffer, arguing that if the polygraph were good enough for the military police, it should be good enough for the court-martial, claimed that this rule, Rule 707, violated his Sixth Amendment right to present evidence in his own defense. The U.S. Court of Military Appeals agreed, but the U.S. Supreme Court did not and reversed. The Court, in an opinion written by Justice Thomas, held that the unreliability of the polygraph justified Rule 707, as did the potential for confusion, prejudice, and delay when using the polygraph.

Justice Thomas, joined by only three other justices (and so not creating a precedent), also wrote that even if the polygraph were extremely reliable, it could not be introduced in court, at least in jury trials. This, he said, was because it too greatly undercut "the jury's core function of making credibility determinations in criminal trials."

Scheffer is a useful reminder that lie detection, whether by polygraph, fMRI, or any other technical method, will have to face not only limits on scientific evidence but other concerns. Under Federal Rule of Evidence 403 (and equivalent state rules), the admission of any evidence is subject to the court's determination that its probative value outweighs its costs in prejudice, confusion, or time. Given the possible damning effect on the jury of a fancy high-tech conclusion that a witness is a liar, Rule 403 might well hold back all but the most accurate lie detection. Other rules involving character testimony might also come into play, particularly if a witness wants to introduce lie-detection evidence to prove that he or she is telling the truth. In Canada, for example, polygraph evidence is excluded not because it is unreliable but because it violates an old common law evidentiary rule against "oath helping" (*R. v. Béland*, 2 S.C.R. 398 [1987]). While nonjudicial use of fMRI-based lie detection is almost unregulated, the courtroom use of fMRI-based lie detection will face special difficulties. The judicial system should be the model for the rest of society. We should not allow any uses of fMRI-based (or other neuroscience-based) lie detection until it is proven sufficiently safe and effective.

A TRULY MODEST PROPOSAL: PREMARKET REGULATION OF LIE DETECTION

Effective lie detection could transform society, particularly the legal system. Although fMRI-based lie detection is clearly not ready for nonresearch uses today, I am genuinely agnostic about its value in ten years (or twenty years, or even five years). It seems plausible to me that some patterns of brain activation will prove to be powerfully effective at distinguishing truth from lies, at least in some situations and with some people. (The potential for undetectable countermeasures is responsible for much of my uncertainty about the future power of neuroscience-based lie detection.)

Of course, "transform" does not have a normative direction—society could be transformed in ways good, bad, or (most likely) mixed. Should we develop effective lie detection, we will need to decide how, and under what circumstances, we want it to be usable, in effect rethinking EPPA in hundreds of nonemployment settings. And we will need to consider how our constitutional rights do and should constrain the use of lie detection. This kind of thinking and regulation will be essential to maximizing the benefits and minimizing the harms of effective lie detection.

But ineffective lie detection has only harms, mitigated by no benefits. As a first step, before consideration of particular uses, we should forbid the nonresearch use of unproven neuroscience-based lie detection. (Similarly, lie detection not based on neuroscience, including specifically the polygraph, should be forbidden. However, polygraphy is likely too well established to make its uprooting politically feasible.)

This step is not radical. We require that new drugs, biologics, and medical devices be proven "safe and effective" to the satisfaction of the federal Food and Drug Administration before they may legally be used outside of (regulated) research. Just as unsafe or ineffective drugs can damage bodies, unsafe or ineffective lie detection can damage lives—the lives of those unjustly treated as a result of inaccurate tests as well as the lives of those harmed because a real villain passed the lie detector. Our society can allow false, exaggerated, or misleading claims and implied claims for many marketed products or services, from brands of beer to used cars to "star naming" companies, because the products, though they may not do much good, are unlikely to do much harm. Lie detection is not benign, but is, instead, potentially quite dangerous and should be regulated as such.

Of course, just calling for regulation through premarket approval leaves a host of questions unsettled. What level of government should regulate these tests? What agency should do the assessments of safety and efficacy? How would one define safety or efficacy in this context? Should we require the equivalent of clinical trials and, if so, with how many of what kinds of people? How effective is effective enough? Should lie-detection companies, or fMRI-based lie-detection examiners, be licensed? And who will pay for all this testing and regulation? As always, the devil is truly in the details.

I have thoughts on the answers to all those questions (see Greely and Illes 2007), based largely on the Food and Drug Administration, but this is not the place to go into them. The important point is the need for some kind of premarket approval process to keep out unproven lie-detection technologies. Thanks to No Lie MRI and Cephos, the time to develop such a regulatory process is yesterday.

CONCLUSION

Lie detection is just one of the many ways in which the revolution in neuroscience seems likely to change our world. Nothing is as important to us, as humans, as our brains. Further and more-detailed knowledge about how those brains work—properly and improperly—is coming and will necessarily change our medicine, our law, our families, and our day-to-day lives. We cannot anticipate all the benefits or all the risks this revolution will bring us, but we can be alert for examples as—or, better, just before—they arise and then do our best to use them in ways that will make our world better, not worse.

Neuroscience-based lie detection could be our first test. If it works, it will force us to rethink a host of questions, mainly revolving around privacy. But if it does not work, or until it is proven to work, it still poses challenges—challenges we must accept. Premarket approval regulation of neuroscience-based lie detection would be a good start.

REFERENCES

Greely, H. T. 2005. Premarket approval regulation for lie detection: An idea whose time may be coming. *American Journal of Bioethics* 5 (2): 50–52.

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

National Academy of Sciences. 2003. *The polygraph and lie detection*. Washington, DC: National Research Council.

Wolpe, P. R., K. R. Foster, and D. D. Langleben. 2005. Emerging neurotechnologies for lie-detection: Promises and perils. *American Journal of Bioethics* 5 (2): 39–49.

Emilio Bizzi is President of the American Academy of Arts and Sciences and Institute Professor at the Massachusetts Institute of Technology, where he has been a member of the faculty since 1969. He is a neuroscientist whose research focuses on movement control and the neural substrate for motor learning. He is also a Fellow of the American Academy of Arts and Sciences, a Fellow of the National Academy of Sciences, a member of the Institute of Medicine, and a member of the Accademia Nazionale dei Lincei. He was awarded the President of Italy's Gold Medal for Scientific Contributions.

Henry T. Greely is the Deane F. and Kate Edelman Johnson Professor of Law and Professor, by courtesy, of Genetics at Stanford University. He specializes in legal and social issues arising from advances in the biosciences. He chairs the California Advisory Committee on Human Stem Cell Research and directs the Stanford Center for Law and the Biosciences. He is a member of the executive committee of the Neuroethics Society and is a co-director of the Law and Neuroscience Project. He graduated from Stanford University in 1974 and from Yale Law School in 1977. He served as a law clerk for Judge John Minor Wisdom of the United States Court of Appeals and for Justice Potter Stewart of the United States Supreme Court. He began teaching at Stanford in 1985.

Steven E. Hyman is Provost of Harvard University and Professor of Neurobiology at Harvard Medical School. From 1996 to 2001, he served as Director of the National Institute of Mental Health (NIMH), the component of the U.S. National Institutes of Health charged with generating the knowledge needed to understand and treat mental illness. He is a member of the Institute of Medicine and a Fellow of the American Academy of Arts and Sciences. He is Editor of the *Annual Review of Neuroscience*.

Nancy Kanwisher is the Ellen Swallow Richard Professor in the Department of Brain & Cognitive Sciences at the Massachusetts Institute of Technology (MIT), and Investigator at MIT's McGovern Institute for Brain Research. She held a MacArthur Fellowship in Peace and International Security after receiving her Ph.D. She then served for several years as a faculty member of the psychology departments at the University of California, Los Angeles and Harvard University. Her research concerns the cognitive and neural mechanisms underlying visual experience, using fMRI and other methods. She received a Troland Research Award from the National Academy of Sciences in 1999, a MacVicar Faculty Fellow Teaching Award from MIT in 2002, and the Golden Brain Award from the Minerva Foundation in 2007. She was elected a member of the National Academy of Sciences in 2005, and a Fellow of the American Academy of Arts and Sciences in 2009.

Stephen J. Morse is Ferdinand Wakeman Hubbell Professor of Law and Professor of Psychology and Law in Psychiatry at the University of Pennsylvania. Trained in both law and psychology at Harvard, he is an expert in criminal and mental health law. His work emphasizes individual responsibility and the relation of the behavioral and neurosciences to responsibility and social control. He is currently Legal Coordinator of the MacArthur Foundation Law and Neuroscience Project and he co-directs the Project's Research Network on Criminal Responsibility and Prediction. He is currently working on a book, *Desert and Disease: Responsibility and Social Control.* He is a founding director of the Neuroethics Society, and prior to joining the Penn faculty, he was the Orrin B. Evans Professor of Law, Psychiatry and the Behavioral Sciences at the University of Southern California.

Elizabeth A. Phelps received her Ph.D. from Princeton University in 1989, served on the faculty of Yale University until 1999, and is currently the Silver Professor of Psychology and Neural Science at New York University. Her laboratory has earned widespread acclaim for its groundbreaking research on how the human brain processes emotion, particularly as it relates to learning, memory, and decision making. Dr. Phelps is the recipient of the 21st Century Scientist Award from the James S. McDonnell Foundation and a Fellow of the American Association for the Advancement of Science and the Society for Experimental Psychology. She has served on the Board of Directors of the Association for Psychological Science and the Society for Neuroethics, was the President of the Society for Neuroeconomics, and is the current editor of the APA journal *Emotion*.

Marcus E. Raichle, a neurologist, is Professor of Radiology and Neurology at Washington University in St. Louis. He heads a pioneering team investigating brain function using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) to map the functional organization of the human brain in health and disease. He joined the faculty of Washington University in 1971, ascending to the rank of professor in 1979. He received a bachelors and a medical degree from University of Washington in Seattle. His many honors include election to the Institute of Medicine in 1991, to the National Academy of Sciences in 1996, and to the American Academy of Arts and Sciences in 1998.

Jed S. Rakoff is a United States District Judge for the Southern District of New York. He also serves on the Governance Board of the MacArthur Foundation Project on Law and Neuroscience, and on the National Academies' Committee to Prepare the Third Edition of the Federal Judges' Manual on Scientific Evidence. He has a B.A. from Swarthmore College, an M.Phil. from Oxford University, and a J.D. from Harvard Law School.

Walter Sinnott-Armstrong is Professor of Philosophy and Hardy Professor of Legal Studies at Dartmouth College, where he has taught since 1981 after receiving a B.A. from Amherst College and a Ph.D. from Yale University. He is currently Vice Chair of the Board of Officers of the American Philosophical Association and Co-Director of the MacArthur Law and Neuroscience Program. He has published extensively on ethics (theoretical and applied), philosophy of law, epistemology, philosophy of religion, and informal logic. His current research focuses on empirical moral psychology as well as law and neuroscience.

The Academy was founded during the American Revolution by John Adams, James Bowdoin, John Hancock, and other leaders who contributed prominently to the establishment of the new nation, its government, and its Constitution. Its purpose was to provide a forum for a select group of scholars, members of the learned professions, and government and business leaders to work together on behalf of the democratic interests of the republic. In the words of the Academy's Charter, enacted in 1780, the "end and design of the institution is . . . to cultivate every art and science which may tend to advance the interest, honour, dignity, and happiness of a free, independent, and virtuous people." Today the Academy is both an honorary learned society and an independent policy research center that conducts multidisciplinary studies of complex and emerging problems. Current Academy research focuses on science and global security; social policy; the humanities and culture; and education. The Academy supports young scholars through its Visiting Scholars Program and Hellman Fellowships in Science and Technology Policy, providing year-long residencies at its Cambridge, Massachusetts, headquarters. The Academy's work is advanced by its 4,600 elected members, who are leaders in the academic disciplines, the arts, business, and public affairs from around the world.