SCIENTIFIC EVIDENCE AND THE CRIMINAL LAW: LESSONS FROM BRAIN-BASED LIE DETECTION

JOHN DANAHER*

I. SCIENTIFIC EVIDENCE AND THE JUDICIAL GATEKEEPER

Science is a powerful tool with many potential forensic applications. Archimedes provides an illustrative example.[1] Archimedes was the most famous scientist of his age, notable for his practical applications of theoretical knowledge. Hiero, the king of Syracuse, enlisted Archimedes's help in resolving his own quasi-legal dilemma. He believed he had been duped by the goldsmith who had made his crown. Hiero had asked for the crown to be made of pure gold, but suspected that silver and copper had been added in order to increase the goldsmith's profit margin.

Archimedes knew that silver and copper were lighter than gold, and would take up more space (*i.e.* have a greater volume) than an equal weight of gold. So in order to determine whether copper and silver had been added, Archimedes simply needed to work out the volume of the crown. This was easier said than done, since the most obvious way to work out the volume would be to melt the crown. This was not an option. How could the matter be resolved?

---

* B.C.L. (U.C.C.), LL.M. (Dub.), Ph.D. Candidate and I.R.C.H.S.S. Scholar, University College Cork. The author would like to thank the Irish Research Council for the Humanities and Social Sciences for their support of his research. Comments and inquiries can be directed to: johndanaher1984 @gmail.com

[1] There seems no doubting that Archimedes discovered the principles of hydrostatics. Whether this particular story about Hiero and the goldsmith is true is a separate question. One is always slightly sceptical of historical anecdotes of this sort. I originally encountered the story in Asimov, I., *Breakthroughs in Science* (Scholastic Magazines, 1959). However, there is reason to doubt the story in its traditional form. The best discussion I have found is on the Archimedes webpage run by Chris Rorres. He points out that a different technique may have been used than that suggested above. See http://www.math.nyu.edu/~crorres/Archimedes/contents.html.

Legend has it that a bath provided the solution. As he lowered himself into the water, he noticed how some liquid spilled-out over the edge, a phenomenon known as displacement. He realised that the volume of water that was displaced was equal to the volume of his body. This was a general principle that could be put to use in ascertaining whether the crown was made of pure gold or not. Using the principle he had discovered, he determined that copper and silver had indeed been added to the gold in the crown. Hiero executed the goldsmith.

We see here how a general scientific law or principle can be put to use in determining someone's legal fate. But questions abound: What if Archimedes got the principle wrong? Would it not then be terrible that a man had died because of mistaken fealty to Archimedes's insight? Is it not incumbent on us to ensure that the scientific evidence used in making legal decisions meets certain epistemic standards? How can we ensure that this is done? These questions lie at the heart of this article.

The article is written against the backdrop of proposed changes to the law in both Ireland and England with respect to the admissibility of scientific evidence.[2] Law reform bodies in both jurisdictions are currently seeking to enhance the gatekeeping role played by the judge when it comes to such admission. This would require a greater critical engagement with the nuances, complications and limitations of scientific inquiry.[3]

The enhancement of the judicial gatekeeping function is something of which I am broadly supportive. But it has to be approached with the right attitude and the right tools. This article aims to provide both. In terms of attitude, it argues that critical engagement with the nature of scientific inquiry is essential, not just in aiding the fact-finding mission of the court, but in

[2] See UK Law Commission, *The Admissibility of Expert Evidence in Criminal Proceedings in England and Wales* (Law Com, CP No. 190, April 2009) hereinafter "Admissibility of Expert Evidence", available at www.lawcom.gov.uk/docs/cp190.pdf. For the Irish equivalent, see *Expert Evidence* (LRC CP 52, December 2008), hereinafter "Expert Evidence". At the time of writing, final reports were pending.

[3] I will provide a brief outline of the proposed tests later in this article, the relevant sections of the consultation papers are as follows: *Admissibility of Expert Evidence* (previous note), part 6; and *Expert Evidence* (previous note), chapter 2, section G.

maintaining the integrity of a liberal democratic system. In terms of tools, it aims to illustrate the basic theoretical and methodological grounding that is needed for critical engagement.

It does so by eschewing an abstract and general discussion of the relevant issues. There will be no comparing and contrasting of the different proposed tests of evidential reliability, nor will there be a general review of the problems courts have had with scientific evidence. These topics have been covered at length elsewhere.[4]

Instead, this article provides an extended analysis of just one type of scientific evidence, namely brain-based lie detection in criminal trials. This practical example will illustrate clearly and forcefully the tools that are required for a critical engagement with scientific evidence and highlight the need for that critical engagement.

Before proceeding with an outline, it is worth pausing to answer the question: why brain-based lie detection? In recent years, increasing attention has been paid to the potential legal applications of the neurosciences.[5] Enhanced tools for investigating, imaging and manipulating the brain have encouraged much academic speculation about neuroethics, neurolaw, neuromarketing and other fanciful neuro-ticisms.[6] But it is not all just idle speculation: there have been several attempts, of varying success, to offer neuroscientific technologies

---

[4] See *Expert Evidence* (note 2 above), chapter 2 for an overview of this topic.

[5] The most conspicuous example of which is probably the MacArthur-funded Law and Neurosciences Project. See http://www.lawneuro.org. Academic articles on the topic are abundant. A representative sampling can be found in Zeki, S. and Goodenough, O., *Law and the Brain* (Oxford: OUP, 2006). Recent conferences in the area include "The Law and Neuroscience Colloquium" University College London, July 6-7 2009; and "Law and Neuroscience: Our Growing Understanding of the Human Brain and Its Impact on our Legal System", organised by the European Science Foundation and held at Hotel Villa del Mare, Acquafredda di Maratea, Italy from 26-30 October 2009.

[6] Examples abound, the setting up of the Neuroethics Society (of which the present author is a member) (see www.neuroethics.com) being conspicuous among them. A neuroethics journal is now published by Springer and the American Journal of Bioethics dedicates three issues per year to neuroethics alone. See Illes, J., *Neuroethics: Defining the Issues in Theory, Practice and Policy* (Oxford: OUP, 2006), for an overview of this academic movement.

for forensic uses. Brain-based lie detection technologies are a good example of this, as will be detailed below.

The remainder of the discussion takes the following form. In section 2, I address some general issues relating to the nature of scientific inquiry. I use this as a springboard for introducing brain-based lie detection, highlighting its likely forms and detailing some of its actual and proposed uses. Section 3 considers how we should critically engage with these technologies. It provides a checklist of questions that any legal practitioner or official should ask about these technologies. It then sketches some possible answers to these questions. Finally, section 4 concludes the article by considering the attitude we should take towards such technologies and toward scientific evidence in general. I make the point that it would be both lazy and dangerous to rely heavily on rights-based objections to these technologies. It is far more important to engage with the details of the science involved, not just for evidential reasons, but for normative reasons as well.

II. THE SCIENCE OF READING MINDS

Science is quite simply our best attempt to understand the world without deceiving ourselves in the process.[7] It is not a body of knowledge per se, but rather a method of inquiry and explanation. Of course, this method has produced a substantial body of knowledge, and familiarity with some of the details of this body of knowledge will help to understanding the cutting-edge of modern science. Nonetheless, we should not lose sight of the fact that we are dealing primarily with a method, not a set of answers.

The scientific method of inquiry is typically taken to proceed in the following manner. First, there exists a set of background assumptions about how the world works. Second, some problem arising from the application of these background assumptions is highlighted. Third, some hypothetical solution to

---

[7] Important since we are quite good at deceiving ourselves. See: Gilovich, T., *How We Know What Isn't So* (New York: The Free Press, 1992) and Sutherland, S., *Irrationality* (2nd Edn, London: Pinter and Martin, 2007) for details.

this problem is proposed. Fourth, this solution is tested in a controlled experimental setting. Fifth, the test either refutes the hypothesis or it does not. And finally, the non-falsified hypothesis is incorporated back into the background theory of the world.[8]

Although there is nothing wrong with this presentation of the method of science (apart from the fact that it idealises the process), I think a better conceptualisation is the following. Science is the method that allows us to best navigate between three universes: the perceived, the detected, and the theoretical.[9]

The perceived universe is that of everyday human sensory experience. The detected universe is that which can be indirectly observed through various detective techniques and technologies. For example, images of human bone can be obtained thanks to our ability to detect and manipulate X-rays. Finally, the theoretical universe is that which explains the perceived and detected universes. For example, the atomic theory in chemistry and physics explains what can be perceived and detected about the many materials with which we interact.

---

[8] This characterisation of scientific method is taken, from Bunge, M. *Social Science Under Debate* (University of Toronto Press, 1998), p. 1. The description therein derives from that author's more ambitious and more informative work *The Philosophy of Science Volumes 1 and 2* (Revised Edn, Transaction Publishers, 1998).

[9] This is taken from the work of David and Richard Garfinkle, as presented in *Three Steps to the Universe* (Chicago: University of Chicago Press, 2008), 1-13. In this book, the Garfinkle brothers look at the means through which scientists build up a picture of the large scale structure of the universe (cosmology). Nonetheless, their idealisation of science is applicable to all other sciences. In discussing their model, a subtle philosophical shift is being made: from scientific methods of inquiry to scientific methods of explanation. I do not mention it in the body of the article, because it would open a philosophical can of worms. Roughly, in discussing method of inquiry we are addressing the procedure of scientific inquiry; in discussing explanation we are addressing what it is that makes scientific explanations successful. Since the scientific method is widely recognised as an idealisation – serendipity is an accepted ingredient – it seems clear that it is more important to talk about scientific explanation. For more on the nature of scientific explanation see Woodward, J., "Scientific Explanation" (2009) in the *Stanford Online Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/scientific-explanation/. Also, Lipton, P., *Inference to Best Explanation* (2nd Edn: Oxford: OUP, 2004). For material that relates specifically to the brain sciences, see Craver, C., *Explaining the Brain* (Oxford: OUP, 2007).

Scientific inquiry proceeds in two directions: from the perceived universe to the theoretical, and from the theoretical to the perceived. What we perceive and detect leads us to question our theories, and the accuracy of our theories is usually tested via the prediction of certain detections and perceptions.[10]

To this three-tiered model of the universe, we can add a fourth tier, that of practical application. This is the point at which scientific knowledge is used to create new technologies. This fourth tier is of crucial importance in the present article, since we are considering technological applications of scientific knowledge in the criminal law.

We can see this four-tiered model of science at work in the example at the start of this article. In working out the relationship between volume and displacement, Archimedes was able to connect elements of the perceived and detected universe with elements of the theoretical universe. He then took this knowledge and put it to use in practical endeavour: determining the guilt or innocence of the goldsmith. As follows: the measurable discrepancies between the weight of different objects, and the volume of water they displace, are part of the perceived and detected universes; while the principle of displacement and the properties of weight and volume belong to the theoretical universe. The principle of displacement was proposed as an explanation of the observations and was then put to use in resolving the practical matter of the goldsmith's honesty.

This four-tiered model will help us to understand the nature and limitations of brain-based lie detection.

---

[10] The classic work on the role of prediction and testability (or, more correctly, falsifiability) is that of Karl Popper. See his *The Logic of Scientific Discovery* (New Edn, London: Routledge Classics, 2002) and *Conjectures and Refutations* (New Edn, London: Routledge Classics, 2002). Popper's theory almost certainly overstates the importance of prediction in the confirmation of scientific theories. For a more up-to-date look at the role that evidence plays in the confirmation of scientific theories, see Sober, E., *Evidence and Evolution* (Cambridge: Cambridge University Press, 2008), chapter 1 "Evidence", as well as his earlier work "Testability" (1999) 73 *Proceedings and Addresses of the American Philosophical Association* 47.

A. Brain-based Lie Detection: The Very Idea

Lie detection is a form of mind-reading. The idea of mind-reading might strike us as exotic, as something that only those with extra-sensory perception or other paranormal powers could engage in. In fact, the practice is far more mundane.[11] Philosophers, psychologists and other behavioural scientists will frequently point out that every human being employs a theory of mind when communicating and interacting with other human beings.[12] We use this theory of mind to explain and predict what our social peers get up to and to structure our responses to their behaviour.[13]

The result is that mind-reading is an almost banal occurrence. That it is banal is highlighted when one considers the devastating repercussions for those that cannot read the minds. One popular interpretation of autism is that it is a deficiency in mind-reading capacity. This causes sufferers to withdraw from social contact.[14]

If mind-reading is so banal, an obvious question arises: what do technologically sophisticated versions add to the mix? Well, there is an important feature to what I will call *mundane* mind-reading that makes alternative mind-reading techniques

---

[11] On these mystics and psychics, although it is hardly a reputable academic publication, Ian Rowland's *The Cold Hard Facts of Cold Reading* (4th Edn, Ian Rowland Ltd, 2002) is recognised by magicians to be one of the best guides to the practice reading minds (or at least giving the illusion of reading minds). It is clear from Rowland's presentation that much so-called psychic mind-reading is really a psychological persuasion technique: subjects become convinced that their minds are being read. This has some disturbing implications for cognitive autonomy.

[12] An excellent exploration of this idea can be found in Tomasello, M. et al, "Understanding and Sharing Intentions: The Origins of Cultural Cognition" (2005) 28 *Behavioral and Brain Sciences* 675. Tomasello et al look at experiments that contrast the human ability to read minds with the chimpanzee ability to read minds.

[13] And also the behaviour of animals and computers. On this point, see McFarland, D., *Guilty Robots, Happy Dogs* (Oxford: OUP, 2008), in particular, chapter 3, "Interpreting Behaviour".

[14] On this see Baron-Cohen, S., *Mindblindness* (Cambridge, MA: MIT Press, 1995). Baron-Cohen has also, along with two others, produced an educational DVD called *Mind-Reading: Teaching Emotion Recognition to those with Autism Spectrum Conditions* (Jessica Kingsley Publishers, 2003).

tempting, namely: it can be easily deceived. This arises from the fact that mundane mind-reading relies entirely on outwardly perceptible cues. In other words, it relies on what people say and do. And of course people can say and do things that are designed to deceive us about their true state of mind. This is particularly so if they have a lot to lose by being honest, as is the case when they are accused of a crime.[15]

Mind-reading technologies try to bypass these potentially deceptive outwardly perceptible cues. Instead, they will use indirect physiological measures, such as the patterns of electrical or vascular activity in the brain, as indicators of true mental state. These physiological measures are thought to be more reliable and less open to deception.

A classic example, of course, is the polygraph lie detector test.[16] The polygraph test measures different levels of activity in the autonomic nervous system[17] (*e.g.* heart rate, and galvanic skin response)[18]. In a typical testing procedure, a subject will be asked a series of control questions, irrelevant questions, and relevant questions. The idea is that measuring the variation in autonomic activity from a baseline established through answers to control questions will indicate whether the subject was lying when responding to the relevant questions.

I will say more about such testing procedures and their reliability later on. For now, I want to draw attention to how the

---

[15] Unintentionally deceptive behaviour can arise from cognitive dissonance. See Tavris, C. and Aronson, E., *Mistakes Were Made (But not by me): How we Justify Foolish Beliefs, Bad Decisions and Hurtful Acts* (Florida: Harcourt, 2007).

[16] Information on how polygraph lie detector tests work is taken from the following sources: National Academies of the Sciences Report, *The Polygraph and Lie Detection* (Washington, DC: National Academies Press, 2003), in particular chapter 3 thereof; Furedy, J., "The North American Polygraph and Psychophysiology: Disinterested, Uninterested and Interested Perspectives" (1996) 21 *International Journal of Psychophysiology* 97; and Furedy, J., and Heselgrave, R., "The Validity of the Lie Detector: A Psychophysiological Perspective" (1988) 15 *Criminal Justice and Behavior* 219.

[17] To be more precise, it measures activity in the sympathetic nervous system, a branch of the autonomic nervous system that prepares the body for responses to stressful situations.

[18] This is a measure of the electrical resistance of the skin. An increased GSR is associated with emotional arousal.

physiological measures are used as indicators of mental state. In mundane mind-reading, the easiest way to gain access to a person's mental state is to simply ask them what they are thinking. If they have no obvious reason deceive us, then this self-reporting will be reasonably accurate.[19] What is more, the answers to these questions will provide us with specific information as to the contents of a person's thoughts. In the case of the polygraph, decoding their mental state is necessarily less precise.

In a polygraph test, the level of physiological activity is taken to be a general indicator of either deception or non-deception. Even at that, the correlation between these physiological activities and deception is indirect, since activity in the autonomic nervous system can be indicative of many things (perhaps most reliably that the person is perceiving an existential threat of some sort). What we have then is a situation in which what we are trying to detect (mental state) is removed from what we actually detect (physiological parameters) by a couple of layers of assumption. Thus any inferences as to likely mental state on the basis of polygraph testing must be made with exceptional care. All mind-reading technologies will suffer from similar indirectness.

The polygraph test remains popular with employers and government agencies the world over, but is of a dubious quality.[20] In the remainder of this article, I want to focus on more recent developments in the area of mind-reading, in particular those arising from brain-scanning technologies such as fMRI (functional magnetic resonance imaging)[21] and EEG (electro-

---

[19] This does not seem to be a baseless assertion. As Michael Tomasello points out, humans, unlike other apes, are incredibly willing to share information. See Tomasello, M., *Why We Cooperate* (Cambridge, MA: MIT Press, 2009).

[20] See *The Polygraph and Lie Detection* (note 16 above) for a comprehensive review.

[21] Most of the information on fMRI is taken from the following sources: Jezzard, Matthews, and Smith, *Functional MRI: An Introduction to Methods* (Oxford: OUP, 2003); Kolb and Wishaw, *An Introduction to Brain and Behavior* (2nd Edn, New York: Worth Publishers, 2006), pp 536-542; Illes, J., Racine, E. and Kirschen, M., "A Picture is Worth a 1000 Words, but which 1000?" in Illes, J. (ed), *Neuroethics* (Oxford: OUP, 2006); Jones et al, "Brain-Imaging for Legal Thinkers: A Guide for the Perplexed" (2009) 5 *Stanford*

encephalography).[22] The former is a measure of different levels of blood oxygenation in the brain. This is taken to be an indirect measure of cognitive activity. Likewise, the EEG measures changes in voltage across the scalp. Since nerve cells communicate with one another through a combination of electrical and chemical signals, the EEG is also taken to be an indirect measure of cognitive activity.

In both cases we have a couple of layers of assumption lying between what is detected and what we really want to detect. The first assumption is that there is a direct correlation between brain activity and mental activity. This assumption rests upon a precarious philosophical peak. Classically, one might have suggested that the mind could not be physical. Although this was always a dubious proposition, it now seems even more unlikely.[23] Data from all the basic and applied sciences suggest that the mind really is linked to physiological processes in the body as a whole and in the brain in particular. But even if we resign ourselves to physicalism about the mind,[24] we cannot simply link particular

---

Technology Law Review (this includes an excellent dissection of an actual research paper – highly recommended); and, more technically, Logothetis, N., "What we can and cannot do with fMRI" (2008) 453 *Nature* 869. There are also many excellent online resources. The best place to go is www.mri-tutorial.com – although this site is incomplete, it has links to many valuable sources.

[22] Information on EEG is taken primarily from Andreassi, J., *Psychophysiology: Human Behavior and Physiological Response* (5th Edn, London: Lawrence Erlbaum, 2007); and Kolb and Wishaw (note 21 above), pp. 140-144. An excellent online resource comes in the shape of John Allen's lectures on the principles of psychophysiology (2004 and 2008 versions). I made considerable use of these in gaining a familiarity with some of the basic aspects of EEG and brainfingerprinting. Available at http://www.u.arizona.edu/~jallen/ (click on "Courses That I Teach" and follow relevant links).

[23] A general review of this debate can be found in Churchland, P., *Matter and Consciousness* (2nd Edn, Cambridge, MA: MIT Press, 1988).

[24] As I happen to think we must. See Melnyk, A., *A Physicalist Manifesto* (Cambridge: Cambridge University Press, 2003) for a defence of this proposition. While I agree with Melnyk, my preferred brand of physicalism is non-reductive in nature. For more on this, see Ladyman, J. and Ross, A., *Everything Must Go: Metaphysics Naturalised* (Oxford: OUP, 2007); and Craver, C., *Explaining the Brain* (Oxford: OUP, 2007).

patterns of neural activity with particular mental states. Not yet at any rate.

The second assumption is that there is a direct correlation between what is measured by these scanners and neural activity. An fMRI detects changes in the oxygenation of blood in the brain. It is assumed (for good reasons) that the nerve cells need oxygen to function properly. Thus the detected changes in oxygenation are thought to reflect neural activity.[25] In the case of an EEG, the measurement of voltage changes may be more direct in that it is clear that nerve cells communicate with electrical signals.

Before moving on to consider how these scanning technologies might become the basis for a legally useful mind-reading technique, let us just briefly consider how mind-reading could be used in criminal law. There are two pairs of possibilities.

The first pair relates to the temporal direction of the assistance. We can either use mind-reading retrospectively or prospectively. Retrospectively, we might use it to try to find out what happened in the past by decoding the information that is present in someone's mind. For example, when a murder has taken place and we want to establish whether someone was present at the scene. Prospectively, we might use these technologies in crime prevention. In this scenario, we might predict that someone is going to engage in criminal activity based on some information present in their mind. On the basis of this prediction, we could take steps to prevent the person from carrying out the predicted course of action. This article limits its purview to the retrospective uses of mind-reading.

The other pair of possibilities relates to the inculpatory or exculpatory uses of mind-reading. Very simply: we can either use mind-reading to link a person to a crime, or not.

---

[25] The connection between neural activity and blood oxygenation now appears to be well understood. This point was made to me by Professor Geoff Aguirre at the University of Pennsylvania Neuroscience Boot Camp, which I attended from Aug 1-12 2009. Professor Aguirre's lectures on the fundamentals of fMRI are available for free download and viewing through his webpage: http://www.cfn.upenn.edu/fundamentals_fmri.htm. See also, Logothetis, N., "What we can and cannot do with fMRI" (note 21 above).

B. MRI Mind-Reading

The popular media is dominated by images of the brain that come from the technique known as magnetic resonance imaging (MRI). MRI scanners can generate static images of brain anatomy and dynamic images of the brain function. The latter come from the technique known as functional MRI (fMRI).

I will now briefly outline the details of MRI and fMRI and then explain how these imaging techniques can be used to read minds in a forensically useful manner. The background details are important because they tie into the overall goal of this article: I am trying to give legal practitioners the tools needed for critical engagement with these technologies.[26]

For the purposes of this discussion, a highly simplified model of the universe is all that is required. We begin at the level of the human body. The human body is made up of different functional systems, for example the nervous system, the circulatory system, the respiratory system and so on. These systems consist of different organs, such as the brain, the heart, and the lungs. These organs are made up of different tissue-types, which are in turn made up of cells. These cells consist of proteins, carbohydrates and other molecular subunits.

Molecules are combinations of atoms in different ratios. Water is one of the simplest and most abundant molecules in the human body. It also happens to play a central role in MRI. It consists of one oxygen atom and two hydrogen atoms. The atoms consist of positively charged nuclei surrounded by negatively charged electrons. The latter are sometimes said to "orbit" the nuclei in shells. Atomic nuclei can be made of two types of particle: protons and neutrons. The proton has a positive charge; the neutron has no charge.

---

[26] See note 21 above for sources used in the following description. On the basic physics, I also made use of the excellent video series by Magritek Ltd., available at www.magritek.com/videos.html. Magritek make MRI scanners, and have provided these videos as an educational resource. If you struggle to understand the concepts of magnetism, precession, torque and resonance after my admittedly brief summary, I would recommend this series.

The hydrogen component of the water molecule has the simplest structure of all, consisting of just one proton and one electron.[27] This proton has angular momentum, *i.e.* it spins around an axis, much as the wheel on a car rotates about an axle. In doing so, it generates a magnetic field with two magnetic poles. If you were to place these hydrogen atoms in a more powerful magnetic field, the protons within them would tend to align with the field lines of this larger magnetic field. The same principle lies behind the alignment of a compass needle with the earth's magnetic field. One of the coils in an MRI-scanner emits a strong magnetic field and when a human body is placed within it, all its hydrogen protons line up with this field.

The angular momentum of the proton causes it to precess (or wobble) about its magnetic poles. The precession occurs at a certain frequency. MRI works off our ability to "see" this precession frequency. When the precession occurs around an axis that is *parallel* to the stronger magnetic field it is difficult to "see", but when the axis is *perpendicular* to the field it is relatively easy to see. One of key functions of an MRI scanner is to cause the protons to "tip over" into this perpendicular plane, thus making the precession frequency visible. This tipping-over requires the application of a torque (twisting force) to the protons. This torque must have the same frequency as the precession frequency. The application of a torque with identical frequency is known as "resonance". In the case of an MRI scanner, this torque is applied via a radiofrequency pulse.

Once tipped over, the protons will gradually realign with the magnetic field. The time taken for this realignment depends upon the viscosity of the tissue in which the water molecule is located. Careful measurement of the time taken for realignment can thus be used to build an image of the internal structure of the body. Of course, all of this is to speak loosely since the precession frequency is not visible in any standard sense, rather, it is capable of being *detected* with the right equipment. An MRI

[27] There are exceptions to this since the presence of neutrons does not affect atomic identity. I am using the simplest possible form of hydrogen atom for the purposes of illustration.

scanner is the right equipment.[28] Through some sophisticated processing, it can use the signals it detects from the hydrogen atoms to recreate an image of various bodily tissues, most famously the brain.[29]

So far the focus has been on MRI scans, which are static. To build an image of brain activity requires fMRI. This simply adds to standard MRI imaging techniques the fact that oxygenated- and deoxygenated-blood have different magnetic properties. Oxygen is attached to haemoglobin molecules in the bloodstream. Oxygenated blood is required for the performance basic metabolic functions. Nerve cells (neurons) require supplies of oxygenated blood in order to function properly.[30] Therefore, it is thought that if one can track the differential flows of oxygenated blood in the brain, one can map which areas of the brain are active at different times.

In addition to this, by getting subjects to perform specific cognitive tasks whilst they are in the fMRI scanner, one can detect which areas of the brain are responsible for the performance of these cognitive tasks. For example, tasks such as mathematical reasoning, visual perception, language production and so on.

When performing such experiments, it is important to remember that there are constant changes and variations in the oxygenation of the blood in the brain. This is because many areas of the brain are active all the time. Only a certain percentage of that variation is relevant to the task being assessed in the

---

[28] An MRI scanner is a long tubular structure consisting of three coils. The first is a superconducting magnet which generates an exceptionally strong magnetic field. When a human body is placed in such a field, the protons within its hydrogen atoms will align themselves with it. A second coil (the gradient coil) corrects for inhomogeneities in this magnetic field and, by producing fields that vary in three orthogonal dimensions, allows us to map locations of hydrogen nuclei in the three spatial dimensions. The third coil generates a radiofrequency pulse. This radiofrequency pulse causes the protons to tip over, thereby making their precession frequency visible. The radiofrequency coil also detects the precession frequency and the time taken for realignment with the magnetic field.

[29] On the issue of image processing, it is worth reading Jones et al, "Brain Imaging for Legal Thinkers: A Guide for the Perplexed" (note 21 above).

[30] See Logothetis (note 21 above) for details on the relationship between neural activity and bloodflow.

experiment.[31] So constructing the functional map will involve the statistical analysis of the different levels and areas of activity over numerous trials. This is an extremely important observation since several recent papers have criticised the statistical tools used by experimenters in deciding which areas of brain activity are correlated with the experimental task.[32]

With this background in place, we can now consider the potential ways in which fMRI could be used to read the mind, remembering at all times that what is being detected is removed from what we desire to know by a couple of layers of assumptions. First, it could be used in a manner similar to the classic polygraph lie detector; and second, it could be used to read information directly from the brain. We will look at each possibility in turn.

As regards the first possibility, an fMRI lie detector will work on the assumption that different areas of the brain will be active when someone is engaging in deliberate deception as compared to when they are telling the truth. Indeed, several studies along these lines suggest that lying involves greater cognitive effort and so results in greater levels of activation when compared to truth-telling. The exact regions of the brain activated during lying vary from study to study, but a popular theory is that lying requires greater input from the executive regions of the brain, *i.e.* the prefrontal cortex and the anterior cingulate cortex.[33]

---

[31] See Jezzard, Matthews and Smith *Functional MRI: An Introduction to Methods* (note 21 above) for more on experimental design. It would also be worth watching Geoff Aguirre's lectures on experimental design (note 25 above).

[32] The classic and most controversial example being Vul et al, "Puzzlingly High Correlations in fMRI studies of Emotion, Personality, and Social Cognition" (2009) 4 *Perspectives on Psychological Sciences* 274. The same issue of that journal contains responses and counter-responses to this paper. This controversy has been circulating for quite some time on the internet and all the back-and-forth is archived here: http://www.edvul.com/voodoocorr.php.

[33] Some examples of such studies would include: Johnson, K. A., George, M., and Kozel, F.A., "Detecting deception using functional magnetic resonance imaging" *Directions in Psychiatry* (2007); Kozel, F.A., et al, "Detecting deception using functional magnetic resonance imaging" (2005) 58 *Biological Psychiatry* 605; Langleben, D.D. et al, "Telling truth from lie in individual subjects with fast event-related fMRI" (2005) 26 *Human Brain Mapping* 262; and Spence et al., "A Cognitive Neurobiological Account of Deception:

On the back of such studies, at least two US-based companies are now offering fMRI lie detector tests for use in legal forums. They are called Cephos Corp and No Lie MRI, respectively.[34] The admissibility of the results of a lie detection test performed by Cephos Corp was decided upon in a recent US-based case called *US v. Semrau*.[35]

The facts of the case are relatively straightforward. Dr. Lorne Allan Semrau ran two companies that specialised in providing psychiatric services to nursing homes in Tennessee and Mississippi. The services were provided under the government-funded Medicare and Medicaid programmes. Dr. Semrau was charged with attempting to defraud these programmes by submitting false and fraudulent claims for payment. The specific details of the charges are not important for present purposes.[36]

In an effort to bolster his defence against these charges, Dr. Semrau enlisted the help of Dr. Steven Laken of Cephos Corp.[37] In order to support Dr. Semrau's claims of honesty, Dr. Laken administered his own patented fMRI lie-detection test. The result was that Dr. Semrau was found to have not lied in response to questions about attempting to defraud the government programmes.

Judge Tu Pham found that these results were not admissible as evidence on several grounds. Chief among them being the fact that the research in this area was in its infancy;

Evidence from Neuroimaging" (2004) 359 *Philosophical Transactions of the Royal Society Series B* 1775. This final article is reprinted in Seki and Goodenough, *Law and the Brain* (note 5 above).

[34] See http://noliemri.com/index.htm and http://www.cephoscorp.com/, for details on both. For a useful discussion of the legal admissibility of such technologies see the special issue of the *American Jounral of Bioethics*, volume 9, September 2007. In particular, the following commentaries: Downie and Murphy, "Inadmissible, Eh?" (2007) 9 *AJOB* 67, and Johnson, K. et al, "The Neuroscience of Functional Magnetic Resonance Imaging for Deception Detection" (2007) 9 *AJOB* 58.

[35] *U.S. v. Lorne Allan Semrau* 31 May 2010 in the US District Court for the Western District of Tennessee (Eastern Division), Case No. 07-10074 M1/P (hereinafter, *Semrau*). I downloaded the case from http://lawneuro.typepad.com/files/semrau.pdf, and all references are to that version. The decision is on the admissibility of the evidence, not on the final merits of the case

[36] Details can be found in *Semrau* (previous note), pp. 3-5

[37] *Semrau* (note 35 above), p. 10

that there were no reliable error rates associated with the technique; and that it was difficult to extrapolate from the results of laboratory tests performed on experimental subjects with nothing to lose, to real-world subjects with plenty to lose, such as Dr. Semrau. The relevance of these criteria to the admissibility of such evidence will be addressed in Section III of this article.

The *Semrau* case is important because it is the first serious consideration of fMRI lie-detection by a court. That it was not accepted is significant, but not necessarily fatal. Indeed, Pham J.'s decision may give companies such as Cephos Corp guidelines as to what they need to do in order to ensure admission in the future.

Moving away from classic lie-detection techniques, the second, and more exotic, potential use of fMRI is to read specific mental content directly from the fMRI image. This is noticeably different from the lie detection case where no specific mental content is read from the fMRI. The possibility of reading specific mental content has emerged in the recent years and is still, unsurprisingly, rather limited. Nevertheless, some of the reported results are impressive. For example, Miyawaki et al managed to discern the type of image that an experimental subject was looking at while being scanned.[38] They were able to do so solely on the basis of the recorded fMRI data.

It should be noted that the visual stimuli involved in this were exceptionally simple two-dimensional monochromatic shapes and letters, not richly-detailed recollections of crime scenes. Also, the time lag between the presentation of the visual stimulus and the recorded fMRI data was relatively short (seconds), but similar experiments with slightly longer time-lags have been reported.[39] This is relevant given that the forensic application of such techniques would require the ability to read complex memories from patterns of activity in the brain. These memories would be of events that long predated the fMRI

---

[38] Miyawaki et al., "Visual Image Reconstruction from Human Brain Activity using a Combination of Multiscale Local Image Decoders" (2008) 60 *Neuron* 915.

[39] Harrison and Tong, "Decoding reveals the contents of visual working memory in early visual areas" (2009) 458 *Nature* 632; Hassabis et al, "Decoding Neuronal Ensembles in the Human Hippocampus" (2009) 19 *Current Biology* 546.

scan. But there is some evidence suggesting that this may eventually be possible.[40]

C. EEG Mind-Reading

The second major brain-scanning technology with potential forensic applications is electroencephalography or EEG. Again, some basic background on the technology is required.[41]

Nerve cells communicate with one another through a combination of electrical and chemical signals. By placing one electrode on the scalp over the brain and another on a relatively electrically neutral part of the head (*e.g.* the ear lobe), an electrical voltage that varies as a function of time can be detected. This time-variant function is known as a "brainwave". This is what is depicted by an EEG monitor. An EEG will usually be built-up from many electrodes placed over different regions of the scalp, not just two as in the example given. This measures variations in voltage over different spatial regions of the brain.

If one were to be hooked up to an EEG, a constant pattern of electrical activity would be detectable. These patterns come in different forms and depend on what the person being recorded is doing. For example, while alert and active, a beta brainwave is most typically observed. These are waves of low amplitude and

[40] Chadwick et al., "Decoding Individual Episodic Memory Traces" (2010) 20 *Current Biology* 544. In this study, researchers got experimental subjects to watch three different videos, depicting different events. They watched these videos 15 times. They were then placed in an MRI scanner and asked to recall one of the videos. Researchers were able to work out which of the videos was being recalled based solely on fMRI data. This involved the decoding of complex memories, of a type similar to those of a crime scene. One significant limitation of this study was that the scans were taken in the early memory-consolidation phase that arises soon after an event. This phase is reliant on a brain structure known as the hippocampus, and it is from scans of that structure that the researchers were able to decode the recollections of the experimental subjects. Long-term storage of memories is known to rely on other brain areas. Decoding recollections from patterns of activity in these other areas may be more problematic. That said, the Chadwick study is an important first step in that direction.

[41] See note 22 above for sources used in the following summary.

high frequency. Higher amplitude and lower frequency waves are observed when the person is asleep (delta waves).[42]

Although the brain is constantly active, and thus constantly generating brainwaves, the presentation of a discrete stimulus would elicit a brief change in the activity pattern. For example, if you were hooked up to an EEG monitor and I were to stamp on your foot or flash a light in your eyes, there would be a "blip" in the pattern of brainwaves being recorded. This is known as an event related potential (ERP). The ERP rides on the existing waves of activity, and so can be difficult to detect. However, repeated presentations of the stimulus can be averaged, and the ERP can then be extracted from the background noise.

There are different types of ERP, which vary depending on the type of stimulus being presented. One particular ERP is known as the P300. The P300 is elicited in response to recognised, meaningful and rare stimuli.[43] For example, suppose I give you a list of twenty unusual words to memorise (say, Latin names of animal species) on Monday. On Tuesday I ask you to sit in front of a screen on which I display a list of Latin animal names, some of which were on the list I gave to you on Monday, and some of which were not. I then ask you to press a button whenever you see a word you recognise. At the same time, I monitor your brainwaves with an EEG. If done properly, I should find that the P300 ERP arises in response to the presentation of the words I gave you to memorise on Monday, but not to the other words.

The fact that the P300 arises in response to recognised, meaningful and rare stimuli is thought to make it apt for use in testing whether a person has concealed, potentially incriminating information stored in their brain. The test might run along the following lines:[44] a suspect is hooked up to an EEG and is seated

---

[42] Detailed discussion of these patterns can be found in Andreassi (note 22 above), pp. 66-70.

[43] Material on the P300 comes from Andreassi (note 22 above) chs 5 and 6, and Rosenfeld, J., "Brain Fingerprinting: A Critical Analysis " (2005) 4 *Scientific Review of Mental Health Practice* 20, available online at http://www.srmhp.org/0401/brain-fingerprinting.html.

[44] Different experimental tests of the P300 are discussed in Rosenfeld (previous note). The description in the text is modified from the descriptions therein.

in front of a computer screen. The screen displays various types of information (words or images). Some of this information will be connected to a particular crime or crime scene (the targets), some will be similar in type but not connected to the particular crime (the probes), and some will be irrelevant. The subject will be asked to press a button in response to recognised information and a P300 should be discernible when responding to recognised information. If the P300 arises in response to the targets, we could infer that the suspect has knowledge linking them to the crime. This would be a guilty knowledge test. It would still involve a form of lie detection. After all, if the suspect admits that they recognise the targets (through the relevant button-pressing exercise), we do not really need the EEG. If, however, they claim not to recognise the targets, but we still observe a P300 response, we can infer that they have the knowledge and are trying to deceive us.

The P300 is the basis of a proprietary technology called brain fingerprinting, developed by the American scientist Lawrence Farwell.[45] Farwell has been offering this technology for a variety of forensic uses for the past 15 years or so. He has met with limited success. He claims that his test can prove scientifically whether a suspect has information linking them to a crime stored in their brains. Such proof would come from a guilty knowledge test such as that outlined above.

It should be noted that Farwell claims his test is not based solely on the P300 response but on something called the MERMER (Memory and Encoding Related Multifaceted Electro-encephalographic Response). The P300 is simply one component of the MERMER. Unsurprisingly, Farwell argues that the MERMER is a far more reliable indicator of guilty knowledge than the P300 alone. However, Rosenfeld, one of the leading electrophysiologists studying the P300, in a lengthy critique of Farwell's technology, has suggested there is no difference between the MERMER and the P300.[46] In any event, the technique for measuring the MERMER has not been released for

[45] Farwell's various claims and publicity materials are available on his website www.brainwavescience.com.

[46] See Rosenfeld (note 43 above) pp. 26-28.

independent confirmation by other investigators. The failure to do so damages the credibility of the technology.

As mentioned above, Farwell's efforts to have brain fingerprinting accepted for forensic uses have met with limited success. He did present evidence in *Harrington v. Iowa*,[47] a case involving a death row inmate. Farwell administered the brain fingerprinting test to Mr. Harrington, which he (Farwell) claimed indicated that Harrington did not have information in his brain that would have linked him to a crime that took place nearly twenty years prior to the administration of the test. Contrariwise, Farwell claimed that Harrington did have information in his brain that linked him to his alibi. Harrington managed to win an appeal to have his case reheard.

There are problems with the retrospective use of the brain fingerprinting test in this case that I will discuss below. For now, all that should be noted is that although this decision is mentioned in the promotional materials on Farwell's website, if one takes the time to read it one will notice that the court felt it unnecessary to consider the reliability of Farwell's evidence in reaching their decision.

In another death row appeal, *Slaughter v. Oklahoma*,[48] Farwell presented evidence suggesting that the convicted murderer did not have specific information stored in his brain that would have placed him at the scene of the crime. The court were unimpressed by the evidence offered, stating:[49]

> Regarding brain fingerprinting, we disposed of that claim in our second post-conviction opinion by noting Petitioner never provided the "comprehensive report" regarding the nature of the brain fingerprinting test conducted, the manner in which it was administered, and the results. We thus found the claim was not backed up with sufficient information for us to act upon.

---

[47] *Harrington v. Iowa*, Supreme Court of Iowa (2003) 659 N.W.2d 509 (No. 122/01-0653, 26 February 2003). Part of the case is excerpted in Rosenfeld (note 43 above), appendix 1.

[48] *Slaughter v. Oklahoma* (2005) OK CR 6, 108 P.3d 1052.

[49] *Ibid*, paragraph 8 of the decision.

On the basis of these two legal airings, it seems unlikely that Farwell's technology will have great success in the forensic arena.

This does not mean that all P300 lie detection techniques or other EEG-based mind-reading tests are forensically defunct. A test developed by an Indian scientist, based on similar principles, has been employed by several police departments in India. It was also, infamously, used to support the conviction of a young woman named Aditi Sharma for the murder of her ex-fiancé in June 2008.[50] As with the *Harrington* case it should not be simply assumed that the mind-reading test was the crucial determinant in that case. Some of the circumstantial evidence considered by the court was, arguably, quite incriminating.[51] In any event, the woman was subsequently released on the grounds that the evidence linking her to the crime was insufficient.[52]

The test in question is known as the BEOS (Brain Electrical Oscillation Signal) test, a name that is singularly uninformative since that is simply a description of what an EEG measures. Despite this, it is claimed, that the BEOS test, using a similar testing paradigm to that outlined above for the P300, can discriminate between *experiential* knowledge and *semantic* knowledge that is present in a suspect's brain.[53] In other words, it can tell whether someone is responding to target stimulus because they have first-hand experience of that stimulus, as opposed to

---

[50] Giridharadas, A., "India's Use of Brain Scans in Courts Dismays Critics" *International Herald Tribune*, 15 September 2008; Natu, N., "This Brain Test Maps the Truth" *The Times of India*, 21 July 2008. The case in question is *State of Maharashtra v. Aditi Baldev Sharma and Pravin Premswarup Khandelwal*, 12 June 2008 (the "Aditi Sharma Case"). The judgment is available at http://lawandbiosciences.wordpress.com/2008/12/10/beos-ruling-from-indian-court/.

[51] The victim was poisoned and the poison was found in some substance in the accused's handbag. That said, the evidence was not well-handled by the police officials and may have been contaminated.

[52] Murphy, E., "Update on the Indian BEOS Case: Accused released on bail", from http://lawandbiosciences.wordpress.com/2009/04/02/update-on-indian-beos-case-accused-released-on-bail/.

[53] I base this largely on the description in the judgment in the Aditi Sharma Case (note 50 above), paragraphs 97-118.

responding because they acquired the target information in some other, presumably non-incriminating, manner.

On the face of it, this would appear to correct for a basic flaw in all guilty knowledge tests: the possibility that someone acquired the "guilty" knowledge in a non-incriminating manner. Furthermore, the idea of a test that can discriminate between different types of knowledge in this manner is not *prima facie* absurd. It is well known that there are different memory systems in the brain, each responsible for storing different types of memory. In fact, there are reported cases involving people who have selective amnesia, following a brain injury, that affects their capacity to recall experiential memories but not semantic ones.[54] And yet, in spite of this, the Indian test lacks credibility for the simple reason that it has not been subjected to independent scrutiny by other scientific researchers.

D. Testing Paradigms

As can be seen from the discussion of both MRI and EEG mind-reading technologies, the information that can be read with the use of brain scans is of limited detail. Both technologies involve deception being inferred from either the activation of different regions of the brain or the presence of an ERP following the presentation of target stimuli.

In order to make reliable inferences to deception, an awareness of the context in which the relevant detection is made is crucial. As noted on several occasions already, the brain is never inactive. We cannot simply place a person in a scanner and expect the relevant information to pop out from the detected patterns of activity. Instead, we must use a series of questions or tasks which make it more likely that the relevant information will be detected. I want to briefly discuss two possible testing paradigms for doing just that: the control question test (CQT) and the guilty knowledge test (GKT).[55] I discussed the latter already

---

[54] Tulving, E., "Episodic Memory: From Mind to Brain" (2002) 53 *Annual Review of Psychology* 1.

[55] Discussion of testing paradigms is based on that found in National Academies of the Sciences Report, *The Polygraph and Lie Detection* (Washington, DC: National Academies Press, 2003), Appendix A, and Furedy, J., and Heselgrave, R, "The Validity of the Lie Detector:

but a little repetition will help to underscore the significance of this issue.

The CQT forms the basis of the classic polygraph lie detector test, but it could easily be used in an fMRI lie detection test. In a CQT a subject is asked a series of questions while their physiological and verbal responses are recorded. The CQT is based on the rationale that a deceptive response to a relevant question will elicit a measurably different physiological response when compared to a truthful response to the same question.

A typical CQT will involve three types of question: relevant, control and irrelevant. Relevant questions are those that are pertinent to the particular crime or event being investigated (*e.g.* did you kill James with a hammer?); control questions are designed to be emotionally similar to relevant questions but unconnected to the particular event (*e.g.* "have you ever lied?"); irrelevant questions are both unconnected to the crime and not of similar emotional weight.

The subject is usually asked to lie in response to the control questions. This establishes a baseline against which the response to the relevant questions will be measured. A subject will be deemed to have lied if the responses to the relevant questions elicit a greater physiological response than the responses to the control questions. In effect, this means that the CQT is a test of the emotional response of a particular person to a particular question. Thus, it becomes a test of the individual, as opposed to a test of any knowledge or information they may have in their brain. A person who fails a CQT can be said to be deceptive, and guilt can, with greater or lesser legitimacy, be inferred from this deceptiveness.

In contrast to the CQT, the GKT does not focus on the person per se, but on the knowledge they may have. In particular, it focuses on guilty knowledge: knowledge that only a person who was present at a crime scene would have. I described this testing paradigm when discussing the P300. Like the CQT, it involves the presentation of three different types of question or

---

A Psychophysiological Perspective" (1988) 15 *Criminal Justice and Behavior* 219.

stimulus: relevant, control and irrelevant – sometimes called "targets", "probes" and "irrelevants", respectively.

On the whole, the GKT is a superior testing paradigm to the CQT. It focuses on knowledge linking a person to a specific crime, whereas the CQT focuses on general deceptive character. Furthermore, the types of control or probe stimuli presented in the GKT can be made to more closely resemble the target stimuli. For example, I could make "did you kill James with a hammer?" the target question, and "did you kill James with an ice-pick?" the control question. The only variation between the questions relates to the murder weapon, with the hammer presumably being the actual murder weapon and therefore the "guilty knowledge". This ability to have limited variation makes the control stimuli in GKT fit the definition of true scientific controls.[56]

In contrast, in the CQT, the control questions can vary greatly from the relevant questions and will tend to be more general. It has been argued, as a result, that the CQT has a tendency to become an interrogation prop. In other words, it is simply a means for encouraging a suspect to confess by convincing them that the test is infallible, or that it has already proven them to be deceptive so they may as well come clean.[57]

It is important to consider the testing paradigm when assessing the validity of the proposed technology. Indeed, I have only scratched the surface of what needs to take place in order to ensure a fair test. One other consideration I would flag would be the need to have formal guidelines for interpreting the results of

---

[56] Controls are used in experimental set-ups to test the validity of a hypothesis. For example, suppose I think that plants grow best when placed in sunlight. To test this hypothesis, I would perform the following experiment. I would place two plant seeds in identical pots, water them with an identical amount at regular intervals and give them an identical amount of plant feed. However, I would place one of the pots in sunlight and the other in a dark room. Thus, the only factor that would vary between the two plants would be the presence of sunlight in one case and its absence in another. The second plant (the one grown in the absence of sunlight) is the control, and it demonstrates that it really was sunlight that helped the plant to grow, not anything else. On this point see Furedy, J., "The North American Polygraph and Psychophysiology: Disinterested, Uninterested and Interested Perspectives" (1996) 21 *International Journal of Psychophysiology* 97.

[57] This is the interpretation of Furedy (previous note), at any rate.

the test and to have two different people administering the test and interpreting its results. This would help to eliminate the possibility of a biased interpretation of the test results.

III. CRITICAL ENGAGEMENT AND RELIABILITY TESTS

The previous section provided some of the necessary background on the relevant brain-scanning technologies and their potential forensic application. This discussion was designed to pave the way for a more substantive consideration of how legal officials can critically engage with scientific evidence of this sort. Critical engagement is exactly what is being demanded by the proposed law reforms in this area. These reforms would require the establishment of formal criteria for the reliability of scientific evidence and judicial training on the application of these criteria.[58]

In this section I am going to detail three questions that any legal official should ask about these technologies. The example of brain-based lie detection maintains its illustrious position as my main reference point, but the discussion is designed to be of more general relevance. At the end of this section, I will compare my three questions with existing and proposed reliability tests.

At the outset, it should be noted that there can be no final or definitive reliability test since the nature of the evidence varies. Nonetheless, useful guidelines are possible.[59]

A. Three Key Questions

To introduce the three questions that form the backbone of my proposed attitude of critical engagement, I want to quickly restate the nature of the evidence with which we are concerned.

---

[58] See *Admissibility of Expert Evidence* (note 2 above), part 6; and *Expert Evidence* (note 2 above), ch 2, Section G, "Junk Science and the Need for a Reliability Test".

[59] See Godden, D., and Walton, D., "Argument from Expert Opinion and Legal Evidence: Critical Questions and Admissibility Criteria of Expert Testimony in the American Legal System" (2006) 19 *Ratio Juris* 261. Walton and Godden provide a reasonably comprehensive set of questions that should be asked of all expert witnesses.

In most aspects of everyday life, we can link a person to an event by simply asking them if they were there at the time (*e.g.* were you at X, at time T? did you do X at time T?). Typically, we will have no reason to suspect dissemblance. However, there is always the potential for deception: a potential that increases when people have a lot to lose in telling the truth. Mind-reading techniques are invariably designed to sidestep the possibility of lying by recording and measuring physiological activity in the brain of person being asked particular questions.

People who advocate the use of these tests will argue that they are valuable aids to the task of criminal justice. In order to assess such value we should ask three questions:

(1) How valid are the theoretical assumptions underlying the technology?

(2) How ecologically valid are the tests?

(3) What is the significance of the result?

Let us consider each of these questions in turn.

1. How valid are the theoretical assumptions?

With this first question we return to the multi-tiered model of the universe outlined earlier. Recall that science is what allows us to successfully navigate between perception, detection and theory. Theory is what explains what we perceive and detect; and the successful prediction of perceptions and detections is what endorses a theory.

In dealing with our fellow human beings we employ a theory of mind. That is to say, we explain what they say and do by positing the existence of certain theoretical entities called beliefs and desires. Proponents of mind-reading technologies accept the validity of this basic theory of mind but argue that the mind is physically instantiated in the brain. So if we can detect the physical state of the brain, then we may be able to detect the state of mind.[60] The question before us is whether these theoretical assumptions are valid. This is undoubtedly a question that is suffused with philosophical import. Indeed, questions about the validity of the theoretical assumptions will always tend to be the most philosophical.

As it happens, I think the theoretical assumptions underlying these technologies are reasonably sound. I think the arguments supporting the idea that the mind is strictly physical are overwhelming.[61] However, some caution is warranted: just because the mind is physical does not necessarily entail that it will be easy to decode specific mental states from patterns of physiological activity.[62] In this regard, the fact that fMRI lie detection tests do not attempt to decode detailed memories, and instead only try to ascertain whether the person is being deceptive, might stand in their favour.

---

[60] Of course, it needs to be remembered that the mind-reading technologies discussed above are only capable of indirectly detecting the physiological activity of the brain.

[61] I would offer three arguments for this being the case: (i) dualism would violate basic physical laws such as that of the conservation of mass and energy; (ii) the materialistic basis of evolution and ontogenetic development; and (iii) the transparent dependency of mental phenomena on neural phenomena as is evinced in cases of brain damage. Some examples of literatures supporting these arguments would include: Churchland, P.M., *The Engine of Reason, The Seat of the Soul* (Cambridge, MA: MIT Press, 1995) and *Neurophilosophy at Work* (Cambridge: Cambridge University Press, 2007); Churchland, P. S., *Neurophilosophy* (Cambridge, MA: MIT Press, 1986) and *Brain-Wise: Studies in Neurophilosophy* (Cambridge, MA: MIT Press, 2004); Bickle, J., *Neuroscience and Philosophy: A Ruthlessly Reductive Account* (Dordrecht: Kluwer, 2003); Damasio, A., *The Feeling of What Happens* (London: Vintage, 1999) and *Looking For Spinoza: Joy, Sorrow and the Feeling Brain* (New York: Harcourt, 2003); Kandel, E., *In Search of Memory: The Emergence of a New Science of Mind* (New York: Norton, 2006); Edelman, G. M. and Tononi, G., *A Universe of Consciousness: How Matter Becomes Imagination* (New York: Basic Books, 2000); LeDoux, J., *Synaptic Self: How Our Brains Become Who we Are* (London: Penguin, 2002); and Llinas, R., *The i of the Vortex: From Neurons to Self* (Cambridge, MA: MIT Press, 2001).

[62] All the works cited in the previous footnote will testify somewhat to this difficulty as they are each a contribution to the literature on the reconciliation problem: how do we reconcile the mental with the physical? However, the research cited in notes 38-40 above suggests that decoding of this sort may become a reality.

But there is an assumption underlying the P300 test that is more worrying. This assumption relates to the nature of memory.[63] It is implied in the guilty knowledge test that the brain is somehow a perfect recorder of an event: that it will store the relevant guilty knowledge in an untainted manner. This assumption is almost certainly false.

There is ample evidence suggesting that memory is a doubly constructive process.[64] What this means is that when memories are initially stored they will be influenced by the knowledge that the individual brings with them to the memorised event. Similarly, when the memory is recalled it will be reconstructed in light of what the person knows at the time of recollection. Furthermore, it is possible to have false memories, *i.e.* memories for events that never took place.[65]

The P300 cannot distinguish between these types of memory,[66] which casts some doubt on its retrospective use, particularly when the event in question took place a long time ago, and there has been plenty of time for the memory to become tainted. Such a situation may have arisen in the *Harrington* case when Farwell's test suggested that Harrington had knowledge of his alibi but did not have knowledge relating to the crime scene. This may have arisen simply because Harrington had ample time to rehearse and ingrain his alibi-story over the elapsed years.[67]

---

[63] This is discussed at length in Rosenfeld (note 43 above), pp. 28-31.

[64] Best overall introduction to this topic is Schacter, D., *The Seven Sins of Memory* (Boston: Houghton Mifflin, 2001).

[65] The became a hugely significant issue in the 1980s and 1990s when many people allegedly "recovered" memories of childhood abuse. It has been argued that many of these cases may have arisen from falsely implanted memories. See: McNally, R., *Remembering Trauma* (Cambridge, MA: Belknap Press, 2003); Pendergrast, M., *Victims of Memory* (London: Harper Collins, 1996); and Loftus, E., and Ketcham, K., *The Myth of Repressed Memory* (New York: St. Martin's Press, 1994) for more.

[66] As was found in Allen, J.J.B. and Mertens, R., "Limitations to the Detection of Deception: True and False recollections are Poorly Distinguished using an Event-related Potential Procedure" (2009) 4 *Social Neuroscience* 473.

[67] In this regard it should be noted that Rosenfeld (note 43 above) argues that Farwell actually misinterprets the results of the test he administered to Harrington. This is based on his reading of the material available on Farwell's website.

Questions about the validity of the theoretical assumptions require a familiarity with the basic science underlying the evidence being presented. Although gaining such a familiarity can be difficult, it is essential if one is to be able to properly assess the evidence in question.

2. How Ecologically Valid is the Experimental Support?

With this second question we turn our attention to the experimental support for the evidence. Whenever possible, scientific evidence should be supported by multiple, independent experimental tests. Indeed, the absence of such tests (as in the case of Farwell's MERMER and the Indian BEOS test) severely damages the credibility of the technique. Such tests will be designed to show that the technique can genuinely distinguish those with guilty knowledge from those without or those who are lying from those who are not.

This will be made possible through careful experimental design in which several subjects are tested and various controls are employed to make sure that what is being detected by the scanner is really due to lying, or the presence of guilty knowledge, and not due to anything else. We should always be careful to check whether proper protocols were followed in these tests. But even if protocols are met we should be keen to question the ecological validity of the experimental design. That is to say, we should ask whether we can legitimately extrapolate from the results achieved in the experimental set-up to the real life scenario with which we are dealing. For example, is it legitimate to extrapolate from an experiment where people are asked to lie about their date of birth or hair colour to a real life scenario in which a person is suspected to be lying about committing a crime? Or is it legitimate to extrapolate from an experiment in which people are given ample opportunity to learn the details of a mock crime scene to a real life crime scene where people may be unlikely to memorise such details?[68]

I do not want to suggest that there is an insuperable bar to such extrapolation – after all, it is never going to be possible to

---

[68] These concerns were raised by Pham J. in the *Semrau* case (note 35 above). He felt the existing experimental tests of the fMRI lie detector did not currently allow extrapolation to real life scenarios.

have a perfect recreation of a real life situation in the lab – but there are some experiments that are going to be closer matches to reality than others. In general, the closer the experiment matches the real life scenario the better. As an aside, it can be observed in the literature on P300 GKTs that the more closely the experimental set-up resembles real crime scenarios the less successful the P300 test of guilty knowledge becomes.[69]

The ecological validity of the test is a question that arises when the scientific evidence in question is supported by controlled experimental tests. In other areas, for example in epidemiological studies, experiments of this nature are off limits. But these studies will still attempt to control for differences between subjects and scenarios, and questions that are equivalent to that of ecological validity can still be asked.[70]

---

[69] A good paper on this point is Mertens, R., and Allen, J., "The role of psychophysiology in forensic assessments: Deception detection, ERPs and virtual reality mock crime scenarios" (2008) 45 *Psychophysiology* 286.

[70] For example, if we want to know whether exposure to mobile phone radiation is linked to cancer, we cannot simply expose some people to the radiation and not others. This would be ethically unsound. What we can do is rely on so-called "natural experiments", *i.e.* compare groups that, as luck would have it, differ only in their exposure to mobile phone radiation. This might involve comparing rates of cancer in previous generations without mobile phones to rates of cancer in the present generation who have mobile phones. Serious questions must be asked about such studies since it will be impossible to control for all confounding variables. If there was a difference between previous generations and the present generation it may be due to differences in lifestyle or diet. Can we rule out those factors? A good, and highly critical, work on this topic is Kabat, G., *Hyping Health Risks: Environmental Hazards in Daily Life and the Science of Epidemiology* (New York: Columbia Press, 2008).

3. What is the Practical Significance of the Evidence?

And so we arrive at the most important question of all.[71] By now we will have accepted both the theoretical assumptions behind the technology and the ecological validity of the experiments supporting it. We will have subjected someone to a test, and we have shown them to be either deceptive or in possession of guilty knowledge. We will be told (based on the experimental support) that the result has certain probability of being accurate. And yet we must still ask what the significance of this result is. There is much to chew over in the process of responding to this query. An illustration of what is at issue will be helpful.

In one of the more infamous trials in recent history, the solicitor Sally Clark was convicted of the murder of her two infant children.[72] The evidence for their murder was scant, and the defence argued that it was possible that they had been victims of SIDS (Sudden Infant Death Syndrome). However, the eminent paediatrician Sir Roy Meadow opined, at trial, that the probability of a double SIDS death in the same family was one in 73 million. Although one can never know what swayed a jury in reaching their guilty verdict, one can assume that Meadow's figures were important. It is now widely known that the figures presented by

[71] There is a line that must be carefully trod here. In this section I am concerned with the practical significance of a test result or a piece of evidence. By "practical" I mean its likely probative value, or bearing on the outcome in a trial. I am not talking about the statistical significance of the test result. This is a technical concept in statistics and experimental design, meaning roughly "what is the probability that an experimentally observed result is due to chance as opposed to being due to whatever it is I am trying to measure?". Statistical significance is a difficult and increasingly problematic concept, see: Ziliak, S., and McCloskey, D., *The Cult of Statistical Significance* (University of Michigan Press, 2008) for a reasonably good, if slightly hysterical, treatment of this topic. A good introductory article would be Siegfried, T., "Odds are, it's Wrong: Science Fails to Face the Shortcomings of Statistics" *Science News*, 27 March 2010. For the purposes of this discussion, all that needs to be known is that a statistically significant result need not be practically significant. If a formal admissibility test is introduced in Ireland, elementary training in statistics would be essential for the judiciary. Rugg, G., *Using Statistics: A Gentle Introduction* (London: McGraw Hill Open University Press, 2007) is a good introductory text.

[72] *R v. Clark* [2000] E.W.C.A. Crim. 54, [2000] All E.R. (D) 1219.

Meadow were erroneous and, indeed, the conviction was later overturned (this later decision relied more on new evidence confirming an alternative cause of death for one of the children and less on the statistical errors).[73]

Meadow in fact made two major statistical errors.[74] First, in arriving at the "one in 73 million" figure, he failed to consider the possibility that shared genetic and environmental factors would make the second death more likely. In other words, he assumed the probabilities for the two events were independent of one another. A more careful treatment of the probabilities suggests that if there has been one SIDS death in a family, the probability of a second SIDS death goes up.[75]

This is, no doubt, a major error, but it is really the second of Meadows's errors that is important to the present discussion. This was his failure to provide the proper context for the one in 73 million figure. It was presented as an isolated figure, with the obvious implication being that there was a one in 73 million chance that she *did not* murder her children. This is an error because, for a proper assessment to be carried out, the likelihood of double SIDS must be weighed against the likelihood of a double murder of the children by their mother. In other words, there were two possible explanations for the deaths on offer and in deciding which was more likely, it was necessary to weigh-up the competing probabilities. There have been post-trial attempts to work out which is more likely, double SIDS or double infanticide, and they suggest double SIDS is more likely.[76]

This is about as clear an example of the importance of asking the significance question as one can find. It illustrates that a positive result on a mind-reading test must not be thought to be the equivalent to a finding of guilt. We must always ask the further question: which is more likely, that the positive result is due to actual guilt or to some other innocent factor. For example, if we are doing a GKT we must ask whether there is a non-

---

[73] *R v. Clark* [2003] E.W.C.A. Crim. 1020.

[74] Detailed treatment of these errors can be found in Hill, R., "Multiple Sudden Infant Deaths – Coincidence or Beyond Coincidence?" (2004) 18 *Pediatric and Perinatal Epidemiology* 320.

[75] *Ibid*, pp. 321-322.

[76] *Ibid*, pp. 322-323.

incriminating reason why the person has the guilty knowledge, and whether the non-incriminating reason is more likely to be true than the incriminating reason. This is particularly important in a system that values the presumption of innocence.

This is tied to another issue bearing on significance, this one concerning the types of error that can be made by the test. Suppose we are told that a particular P300 GKT is 90% accurate. As it stands this figure is insufficiently informative because it does not tell us the ratio of false positives to false negatives. Indeed, a sneaky advocate could make the 10% of failures refer to either the number of false positives, false negatives or both. False positives and false negatives are the two types of error that can be made in a test of this nature. A false positive error would arise when a P300 is elicited from a person without the guilty knowledge; a false negative error would arise when a P300 fails to be elicited from a person with the guilty knowledge.

It would be relatively easy to work out the ratio of false positives to false negatives in an experimental set-up where those with the guilty knowledge can be readily identified. Imagine we do test with 20 people, 10 of whom have been given access to guilty knowledge and 10 of whom have not. We then administer the P300 GKT. The result is that we successfully identify 8 of the 10 with guilty knowledge, but at the same time falsely identify 2 of those without the guilty knowledge as having it. Thus we have 10% false negatives and 10% false positives. The question is whether it is acceptable to falsely accuse those two people and whether this is an issue that should be put before a jury.

Blackstone famously commented that its better that nine guilty men go free than that one innocent man be found guilty. Do we accept this epithet today? If so, we should prefer a test with as low a rate of false positives as possible. How low that rate of false positives needs to be is going to be a bit of a judgement call. Obviously, no test is going to be 100% accurate, but at the same time a test with, say, 25% false positives would seem to be excessive. My guess is that less than 10% would be desirable, provided we still remember not to equate a positive result with a finding of guilt. Interestingly, Blackstone's epithet suggests that a high level of false negatives is acceptable. However, a test with a very high rate of false negatives might become next to useless: its results would have no probative value.

The significance question is designed to make sure we do not jump to conclusions. Even if we have a mind-reading test with sound theoretical assumptions, good ecological validity and high accuracy, we must still consider the significance of a positive result in light of other considerations (error rate and alternative explanations). This is a question that must be asked at both the admissibility phase and the decision-making phase. At the former phase we must ensure that the evidence is reliable enough to be entered into consideration. And at the latter phase, we must ensure that we appreciate the contextual significance of the evidence. If a jury is deciding the question of guilt, they must be reminded of the question.

B. Reliability Tests

The three questions just outlined are the types of question that I think need to be asked when considering the admissibility[77] of brain-based lie detection and cognate scientific evidence. How well do these questions compare with the reliability tests that are proposed for England and Ireland? I would say they compare reasonably well as both proposals would require considerable engagement with the nature and limitations of scientific inquiry.

In their consultation paper on this issue, the UK Law Commission propose a test that would involve, *inter alia*, consideration of the principles underlying the evidence, the testability of those principles, the error rates associated with the evidence, whether the evidence has been accepted in the scientific community, the validity of alternative theories, and the qualifications of the expert.[78] The Irish Law Reform Commission have suggested a nearly identical set of criteria.[79] The Irish proposal also argues that the judge should begin with the assumption that the evidence is unreliable when assessing it.[80]

---

[77] Although, as noted, the final question is important at the decision-making phase as well.
[78] *Admissibility of Expert Evidence* (note 2 above), pp. 53-53.
[79] *Expert Evidence* (note 2 above), pp. 128-129.
[80] *Ibid*.

In this article, I have attempted to show what these proposed tests would entail in practice by using the example of brain-based lie detection. As can be seen, effective critical engagement with these techniques entails quite a lot: it requires some knowledge of the basic science and theoretical assumptions on which they are based; it requires knowledge of the circumstances in which these techniques are tested; and it requires detailed consideration of the error rates associated with them.

IV. CONCLUSION: THE CORRECT ATTITUDE?

I started this article with two goals. First, by using the example of brain-based lie detection, I wished to provide legal practitioners with the tools needed to critically engage with scientific evidence. It is hoped that the previous sections have achieved this goal. But that brings me to the second of my original goals: providing the correct attitude with which to approach scientific evidence of this sort. I will conclude by considering this issue.

An obvious concern arises from the foregoing material: in asking judges (or, indeed, juries) to critically engage with scientific evidence, are we not demanding too much? Life is short and science is increasingly complex, why not outsource the difficult task of dealing with this evidence to those who are most qualified to do so?[81] I think to adopt such an attitude would be a mistake. Science, and the technologies to which it gives rise, is deeply embedded in contemporary life. To leave it in the hands of an elite few, would be a recipe for disaster.

Look at it this way: the thought that there exist technologies that are capable of reading our minds may seem frightening. This is particularly so if we adhere to a certain normative ideal for society. For example, I admire the vision of society that is outlined in John Stuart Mill's *On Liberty*.[82] Therein, Mill makes a powerful argument that society should aim to maximise the flourishing of the individual. Each person should be allowed to engage in creative experiments in living, while

---

[81] For a proposal along these lines see Robertson, C.T., "Blind Expertise" (2010) 85 *New York University Law Review* 174

[82] (London, 1859).

being protected from the inertia of custom, the slumber of reason and the tyranny of both the majority and the magistrate. The only countervailing pressure is that the individual should not be allowed to harm others.

The prospect of mind-reading may seem to be anathema to this vision. If law enforcement officials can tell what you are really thinking it may prevent freedom of thought and expression; it may suppress creativity and individual flourishing; it may enable the tyranny of the magistrate. We might be tempted into thinking that the prudential reaction to such a possibility would be to simply argue for a more robust and stringent protection for individual rights.[83]

I think there is a danger in this type of reaction since it tends to grant the claims made by advocates of the technology a *prima facie* credibility that they do not deserve. The correct attitude towards such developments is the one of critical engagement not reactionary detachment. The legal system should not leave the engagement to the elite few. To illustrate this point, consider that the Royal Society was set up 350 years ago as an institution for the furtherance of scientific inquiry. Its motto was, and still is, "*nullius in verba*" – "on the word of no one".[84] The idea here is that scientific inquiry is the ultimate form public inquiry; that the findings speak for themselves; that they require no faith in the opinions of authorities.

There is a lesson here for modern society which has become so reliant on the fruits of scientific inquiry: If you treat scientific knowledge as something that is abstruse and technically difficult, if you sensationalise the details by failing to engage with them, then you will render yourself open to cognitive exploitation. That would be when, to return to Mill, we would truly lose our liberty.

[83] As for example in Halliburton, C., "Letting *Katz* out of the Bag: Cognitive Freedom and Fourth Amendment Fidelity" (2007) 59 *Hastings Law Journal* 309, specifically discussing brain fingerprinting and brain scanning technologies.
[84] For a discussion, see http://royalsociety.org/Nullius-in-verba/.