ARIZONA STATE UNIVERSITY

SCHOOL OF SUSTAINABLE ENGINEERING AND THE BUILT ENVIRONMENT

CESEM

Center for Earth Systems Engineering and Management

Worrying About Our (Neuro) Image:
How much does fMRI really reveal about us?

Tom Roberts

ASU-SSEBE-CESEM-2011-RPR-002
Research Project Report Series

May 2011

ASU Ira A. Fulton Schools of Engineering

Worrying About Our (Neuro) Image:
How much does fMRI really reveal about us?

Tom Roberts
May 4, 2011

Abstract

After a brief introduction to Functional Magnetic Resonance Imaging (fMRI), this paper presents some common misunderstandings and problems that are frequently overlooked in the application of the technology. Then, in three progressively more involved examples, the paper demonstrates (a) how use of fMRI in pre-surgical mapping shows promise, (b) how its use in lie detection seems questionable, and (c) how employing it in defining personhood is useless and pointless. Finally, in making a case for *emergentism*, the paper concludes that fMRI cannot really tell us as much about ourselves as we had hoped. Since we are more than our brains, even if fMRI were perfect, it is not enough.

Introduction

"I yam what I yam and tha's all I yam" – Popeye.

If you are unfamiliar with Popeye the Sailorman, you've missed a wonderful illustration of the nature v. nurture debate. Popeye is an interesting case because despite his previously quoted protestation to the contrary, he is frequently much more than he seems. When pressed to save his girl—the fickle and flighty Olive Oyl—from his arch nemesis Brutus, Popeye simply pours canned spinach down his throat and instantly processes the magic muck into muscle mass (not to mention some courage and determination), enabling him to perform heroic feats. This merging of nature (his obviously incredible but not well understood genetic constitution) with nurture (the environmental insult of romance, stress, and spinach) demonstrates the complexities of real life—who we are and what we are seem to be difficult questions to answer. It also serves to put the lie to Popeye's favorite expression. As it turns out, as humans, we are what we are—and then some!

Popeye's curious melding of philosophy and hard science constitute a solid demonstration that this research could go one of three ways. On one hand, it could turn into a *philosophical* treatise on how neuroscience is redefining humanity and personhood, and we could spend our time discussing the timeless arguments of many philosophers living and dead. In doing so, we could wrestle with the age-old questions of who we are and if (or to what extent) we are free. Ultimately, we could arrive at a well-fortified determination of whether or not Popeye was right in his self assessment.

On the other hand, this research could be about *altering* the brain using drugs or other mechanisms for enhancing our performance (we were, after all, speaking about Popeye). Drugs that augment cognition are quickly becoming ubiquitous and other protocols such as deep brain stimulation are being introduced regularly and rightly receiving more than the dismissive nod reserved for quacks and scientific sensationalists. There is merit in pursuing such discussion from both the physiological and ethical standpoints. While we may never learn the secrets of Popeye's enhanced capabilities, there is great value in studying the impact of pharmaceuticals on our physiology.

This paper, however, takes neither of these courses. Instead, this paper discusses *monitoring* the brain and *reporting* what is seen—specifically through the technology of Functional Magnetic Resonance Imaging (fMRI). In the developed world, it is difficult to find someone who has not been exposed to, witnessed, or been greatly impacted by the powerful images generated by this sophisticated technology. fMRI is being used increasingly and to great effect in allowing us to look inside our bodies and brains. But, just how much can fMRI tell us about ourselves? Sure, it can show us a tumor, but can it answer any deep questions about who we are, or even how we think? In an ironic twist, our discussion of technology that *looks at* the brain leads us back to more philosophical contemplation on *how we think about our thinker*.

Science is aggressive and will purport to tell us who we are. In fact, fMRI has been able to tell us more about ourselves than any other science to date. Just how much does it tell us? How much can we believe?

Functional Magnetic Resonance Imaging (fMRI)

Functional Magnetic Resonance Imaging (fMRI) has become the poster child of neuroimaging over the past 20 years. Despite incredible progress and promise, it can still be rightly identified as an emerging technology and is hence not quite ready for primetime in many areas of application. Still, with a long lineage of strong technology behind it, fMRI makes a great case for its ability to provide future contributions.

Nuclear Magnetic Resonance is the science that studies how protons in atomic nuclei have a magnetic moment (angular momentum, i.e., spin) that absorbs and releases energy at a specific resonance frequency. NMR science was first exploited by Raymond Damadian in 1971 to detect tumors and by 1977 the first clinical use of an MRI scanner was performed to create the first body scan images. MRI has been successful in generating high resolution "still" images for years and has many clinical applications including diagnoses of tumors. In 1990, Seiji Ogawa observed the Blood Oxygenation Level Dependent (BOLD) effect that is now instrumental in nearly all fMRI use. The BOLD response (discussed later) allows the technicians to observe a body or brain while it is functioning (though it is more like a cartoon flipbook effect at one image every few seconds than it is a motion picture at 30 frames per second). By 1992, Ogawa published the first functional images using BOLD signal, and the fMRI was born. Twenty years later, fMRI is still emerging, and a bit of haiku sets the stage for our look at its mysteries:

> *Magnets dim the lights.*
> *Hydrogen's pole dance begins.*
> *C'est l'amour! Oui? No!*

How it Works

Figure 1 depicts the basic equipment configuration for fMRI. The primary magnet is configured to produce a very strong field centered at the *bore*, where the subject's head is positioned. Since the subject is expected to remain effectively motionless (head motion of more than 2mm can render scans useless), they are sometimes equipped with special glasses that enable them to see a prompt screen that may contain instructions or questions that are part of the experiment. There are frequently buttons provided to enable the subject to record decisions or selections made. Control computers record the stimulus, responses, and the brain function during the experiment.

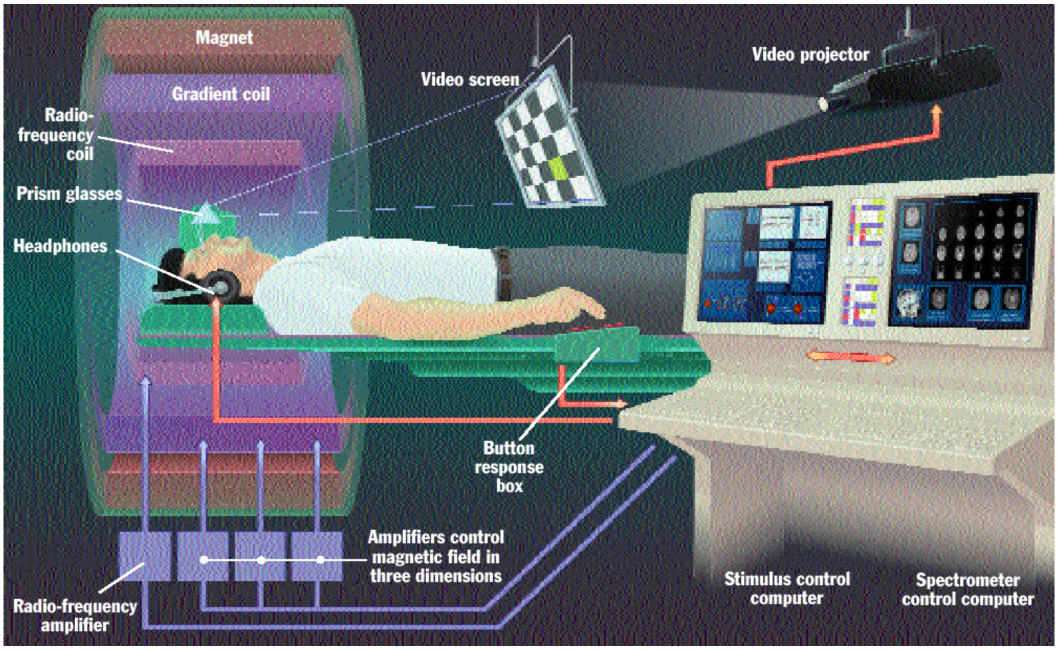

Figure 1. Typical fMRI Equipment Configuration
(Source: Culham, 2008, University of Western Ontario)

fMRI is referred to as "functional" because, unlike its non-functional counterpart that produces high quality static pictures of brain structure, fMRI measures the working brain by taking a sequence of lower resolution "snapshots" as the brain is responding to stimuli. Such "measurement" is based on a series of assumptions about how the brain works.

First, it is generally accepted that neurons do the "thought work" of the brain. But fMRI does not directly measure neuronal activity. Instead, when neurons are working ("firing"), they burn stored energy and transfer chemicals (neurotransmitters) which thereafter must be restored during a recovery period in which the neurons reset. So, the second assumption is that blood flow in the brain restores this energy through a complicated and not yet well understood metabolic pathway involving astrocytes at the so-called "blood-brain barrier" (see Figure 2). It is assumed that an insurgence of oxygenated blood will follow shortly after a burst of neuronal activity—though this is delayed by up to several seconds. In this manner, blood flow is assumed to be indicative of neuronal activity in the brain and is measured as a proxy. (Figley & Stroman, 2011; de Zwart et al., 2005; Matthews & Jezzard, 2004).

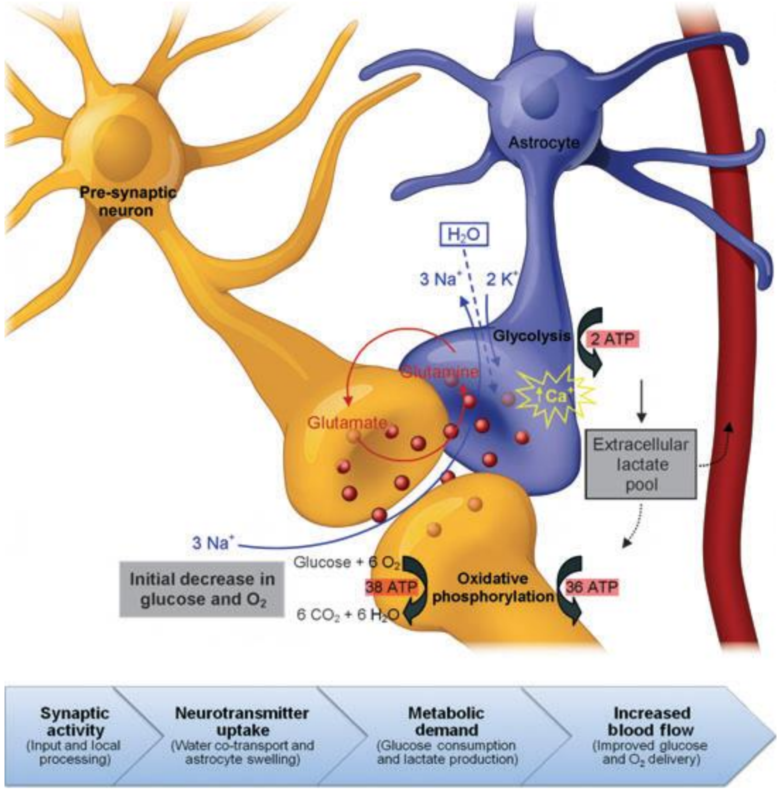

Figure 2. Astrocyte role in neurotransmitter recycling and cerebral blood delivery in the brain
(Source: Figley & Stroman, 2011, *European Journal of Neuroscience*)

The blood flow itself can be measured through a complicated procedure that is made possible by the Magnetic Resonance Imaging equipment. The large, superconducting magnet in which the subject is positioned is used to align a (typically small) percentage (~5%) of the spinning hydrogen nuclei (protons) of the water in the blood. This will tend to induce a bulk magnetization aligned with that of the large magnetic field with a resonant frequency known as the Larmor frequency. The orientation of the protons' magnetic moments can be perturbed (through a radio frequency pulse that matches the precession frequency) such that it is temporarily oriented perpendicular to the magnetic field. This will induce a voltage that can be measured as the protons slowly return to their original orientation. Since the molecules in the blood have different magnetic properties based on the amount of oxygen they contain, the difference between oxygenated blood and deoxygenated blood can be recognized. The result is that a higher signal is detected for oxygenated blood and this is known as Blood Oxygenation Level Dependent (BOLD) contrast imaging (Logothetis, 2001; Logothetis, 2008; Poldrack, 2008, Amaro & Barker, 2006).

Figure 3 depicts the proxy relationship between the blood flow and the neuronal activity and highlights (at step 2) the uncertainty in understanding about the metabolic signal that provides the neurovascular coupling. While our knowledge of this metabolic pathway and its relationship to neuronal activity is improving, it is still somewhat of a mystery and research continues (Figley & Stroman, 2011; Koehler, Roman & Harder, 2009; Pellerin, 2005; Tagamets & Horwitz, 2001).

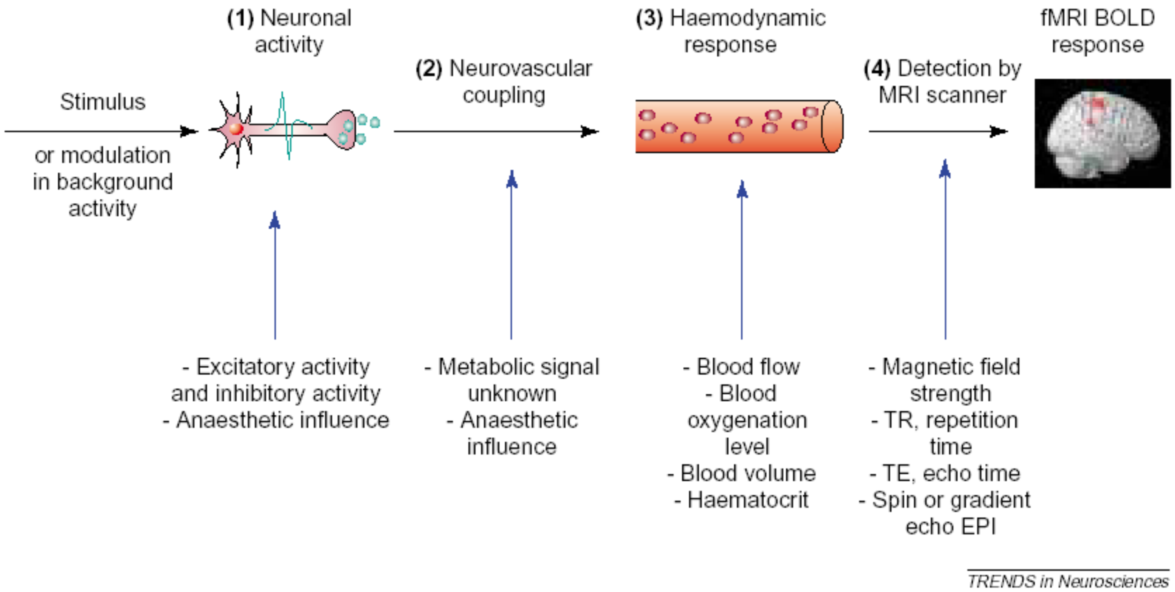

Figure 3. Neuronal Activity and the BOLD Response
(Source: Arthurs & Boniface, 2002, *Trends in Neurosciences*)

The electronics in the fMRI apparatus repeatedly takes snapshots of the subject brain's BOLD hemodynamic response function (HRF) and builds up a collection of images that depicts neuronal activity (through the proxy of the blood) at any given time in the imaging sequence. A third assumption is then made about the response function: that certain areas of the brain are associated with certain kinds of mental processing (e.g., in very general terms, the pre-frontal cortex is associated with executive function and decision making), and that the fMRI has adequate resolution (both temporal and spatial) to detect such processing. A significant percentage of fMRI research has focused on locating what are known as the "neural correlates" for certain kinds of mental processing. While mapping mental function to brain structure is fraught with challenge and constitutes more of an art than a science, there has been some *limited* success in repeating experiments and consistently implicating certain brain regions with certain mental activity (Poldrack, 2010; Poldrack 2006; Devlin & Poldrack, 2006; Van Horn & Poldrack, 2009).

Can we trust fMRI?

To say the least, fMRI technology is exceedingly complicated. But is it so complicated that we can't really trust it? Our trust in the science must hinge on the three assumptions presented above:

1. *If* blood flow is an adequate proxy for neuronal activity, *and*
2. *If* fMRI resolution in measuring the blood flow is adequate in both temporal and spatial dimensions, *and*
3. *If* regions of increased blood flow indicate specific areas in which specific mental processing occurs (i.e., if the concept of "neural correlates of cognition" is true),

then, we can reasonably expect that fMRI is accurately reflecting brain function.

But we must still caveat that conclusion with the idea that while we may know *that* you're thinking, we still don't know *what* you're thinking. Many researchers have needlessly reminded us that while significant progress is being made, fMRI is not (yet?) mind-reading (e.g., Brown & Murphy, 2009, but see Cox, 2003; Norman, 2006).

The first assumption does appear reasonable even given the complexity of the astrocyte interactions and involvement in rejuvenating the synapses after neuronal activity. Whether or not the neurons restore themselves and whether or not neurons are the cells consuming the glucose (as measured in, for example, PET and SPECT modalities), once the neurons "fire" there is something in the region that consumes blood to restore the synapse for its next use (Figley & Stroman, 2011). Despite the complexity and ongoing research, it appears that blood flow is a good indication of neuronal activity. It is however, delayed in time by anywhere from 1 to 5 seconds (de Zwart et al., 2005) and occurs over very broad volumes of the brain (Logothetis, 2008).

It is with the second assumption that things become less clear. Logothetis sets the stage for our discussion of belief in this assumption with his description of what is actually in a volume (think of a *voxel* as a volume-pixel) of brain matter that is measured by fMRI:

> What are the actual contents of a neuroimaging voxel? An examination of the 300 top-cited cognitive fMRI studies suggests that the commonly used in-plane resolution is 9–16 mm^2, for slice thicknesses of 5-7 mm. The average voxel size before any pre-processing of the data is thus 55 ml (or $55mm^3$). Often the effective size is 2-3 times larger due to the spatial filtering that most investigators apply to improve the functional [signal to noise ratio]. Less than 3% of this volume is occupied by vessels and the rest by neural elements. A typical unfiltered fMRI voxel of 55 ml in size thus contains 5.5 million neurons, 2.2–5.5 x 10^{10} synapses, 22 km of dendrites and 220 km of axons (Logothetis, 2008, p. 875).

Based on this background information, it is not an overstatement to say that it would be difficult to tell what's going on in that very complicated space. Even when a significant activation is measured in this little region it is generally accompanied by activity in the surrounding regions and so measurements of such activations must be "smoothed" by after-the-fact image-processing software. Logothetis is being conservative when he suggests "the effective size is 2-3 times larger due to spatial filtering" that occurs during this smoothing process. In fact, Spence et al. published results with an extent threshold of 30 voxels (Spence et al., 2004, p. 1761)! Such activations must also be compared to "steady state" brain activity (sometimes referred to as "default mode") and somehow determined to be statistically significant. There are many ways this is accomplished (Amaro & Barker, 2006; Haller & Bartsch, 2009) the most rudimentary of which is simple subtraction which is almost never used, but is easily illustrated to a novice as Spence (2008) does when discussing his mechanisms for detecting lies with fMRI (more on that topic later):

[Lie scans] - [Truth scans] = Areas of activity associated with lying

Another important fact is that only part of the activity being measured is "excitation" related. Scientists have determined that some of the activity is actually inhibitory in nature (Logothetis, 2008). Finally, the software employed to perform the smoothing and mapping to a "common" brain structure is getting better and more sophisticated, but it is far from standardized (Devlin & Poldrack, 2007; Fadiga, 2007; Toga & Thompson, 2007), leaving much margin for experimenter interpretation.

Regarding the third assumption, most neuroscientists do assume there are regions of the brain that perform certain kinds of processing, but studies that rely on this so-called "reverse inference" to determine "neural correlates" have come under severe scrutiny in the literature (e.g., Poldrack, 2008). Such approaches have become so rampant, that some foundations have simply refused to fund research that makes such assumptions. Fadiga reports, for example, the guidelines of the James S. McDonnell Foundation in their own words:

> Proposals to use functional imaging to identify the “neural correlates” of cognitive or behavioral tasks (for example, mapping the parts of the brain that “light up” when different groups of subjects play chess, solve physics problems, or choose apples over oranges) are not funded through this program. In general, JSMF and its expert advisors have taken an unfavorable view of projects attempting too wide a leap in a single bound. Functional imaging studies using poorly characterized tasks as proxies for complex behavioral issues involving empathy, moral judgments, or social decision-making are generally not appropriate responses to this call for proposals (Fadiga, 2007, p. 1042).

To say the “fMRI industry” has become jaded by thousands of studies that allege to demonstrate trivial notions like, for example, “when viewing chocolate, chocoholics demonstrate an increase in activity in the reward center of the brain” would be an understatement (Beck, 2010). But, as described earlier, it is not just the sponsors who have become suspicious. Researchers are now seriously contending with hard questions about just how far we can go in proving neural correlates for specific kinds of processing (Uttal, 2001, 2002a, 2002b; Poldrack, 2010; Gonsalves & Cohen, 2010; Van Horn & Poldrack, 2009; Devlin & Poldrack, 2007; Toga & Thompson, 2007; Poldrack, 2006; Yarkoni, 2009). William Uttal of Arizona State University may have cast the initial doubts in his 2001 book entitled *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*, which attacks the issue head-on as he describes:

> In short, my argument is that the enormous effort currently being made to localize vaguely defined cognitive modules as attention or language in particular locations of the exceedingly complex brain is an ill-chosen path for cognitive neuroscience. I believe that it is based on incorrect a priori assumptions that, on close examination, cannot be justified and data that is highly tainted by these incorrect assumptions. Such an approach is leading us astray from thinking (correctly, I argue) about cognitive activity being broadly distributed in the brain (Uttal, 2002a, p. 226).

In a humorous turn not often seen in the ranks of formal academia, one research group refers to such mapping of neural correlates as “voodoo” (Vul et al., 2009, p. 274; Diener, 2010). And while such hyperbole has no real scientific value, it does at least remind us that fMRI is still as much of an art as it is a science.

Common Misunderstandings about fMRI

One of the more common misunderstandings of fMRI is that it takes photographs of the brain. While this is a convenient metaphor for non-scientists, it is vital to understand that what fMRI does is significantly different. As Adina Roskies reminds us “although neuroimaging does bear important similarities to photography, the details of the generation and analysis of neuroimages significantly complicate the relation of the image to the data” (Roskies, 2008, p. 19). While this has little to do with the experimental design or measurement protocols, it does have a large impact on how the results are evaluated and judged by an ignorant public. It is difficult to not be impacted by the so-called Christmas tree effect (Mobbs et al., 2007). Hence, it is vital to remember that fMRI measures tiny magnetic field fluctuations and miniscule voltage differentials from a small percentage of hydrogen protons in blood flowing in comparatively large volumes of the brain that is only indirectly related to the neuronal activity that occurred several seconds prior to the measurements being taken. As if that isn’t enough, this data is then massaged and subtracted, statistically analyzed and extrapolated, colorized and mapped onto a generic brain before it is finally rendered for display. Despite their absolute and unequivocal remoteness from reality, the pictures are surprisingly useful.

Additionally, it is often forgotten the extent to which human factors impact the results of an fMRI scan. Issues such as respiration, motion, drugs, age, attention, and a variety of neurovascular coupling differences can greatly alter the quality of the images generated (Haller & Bartsch, 2009). Such things are constantly plaguing researchers and impacting their results. It is not uncommon to read in the peer-reviewed literature phrases to the effect of "14 subjects were scanned and 6 were rejected due to head motion greater than 2 mm." Further, older people are reported to be more distracted by the noise of the gradient magnets as they realign for the progressive image slices, thus impacting scan results due to subject inattentiveness (Stevens et al., 2008). These small things tend to degrade the data and have negative effects on the results. Add to this a host of experimental factors like measuring "activation" versus "baseline" brain activity, on-off experimental paradigms, block v. event designs, etc. (Amaro & Barker, 2006) and the researcher has his hands full wrestling a Lernaean Hydra. Still, sometimes the scientist wins.

One final misunderstanding is the extent to which data processing software is involved in generating and manipulating the results of fMRI scans. Not only is there sophisticated data acquisition software controlling the timing and other parameters of the experiment, but there is an extensive suite of data processing software employed for image alignment, motion correction, smoothing, and noise reduction. To successfully measure signals that are so faint, humans cannot be in the loop—except in the form of the software they've written. Even after the data has been collected, sophisticated data interpretation tools are employed to maximize the experimental results (Haller & Bartsch, 2009; de Zwart et al., 2005; Matthews & Jezzard, 2004). It is well beyond the scope of this paper to explain the details and complexity of such software, but suffice it to say it is far more complicated than developing a photograph.

The Problem with Group Studies

Many scholars have outlined the issues with group studies and fMRI. Many others have invested significant effort in solving the problems and enabling the reporting of credible data. The extraordinarily rich area of literature centered on this topic would be impossible to summarize herein (e.g., Mumford & Poldrack, 2007; Mumford & Nichols, 2009; Stephan et al., 2009; Ramsey et al., 2010). Instead, what is offered is a simplified explanation of the rudiments of the problem.

Figure 4 provides a notional depiction of the approach used to formulate results from group studies. The top row of images contains (first) the colorized image of an "activated" brain performing some action or receiving some stimulus. The second image in the top row represents the same brain with that stimulus removed. The third image shows a simple notional subtraction of the two images leaving what is intended to show the brain's active areas when performing the action or receiving the stimulus of the first image. What has been subtracted can be considered "background" or "default mode" brain function (Raichle et al., 2001; Buckner et al., 2008). The second row of images provides a notional illustration of the subject in the top row (as indicated by the arrow) and four other subjects after the same stimulation and control subtraction had been performed as for the first. Note the differences in the activations across the unique subjects. The single image in the final row of the figure shows the average activation for the group.

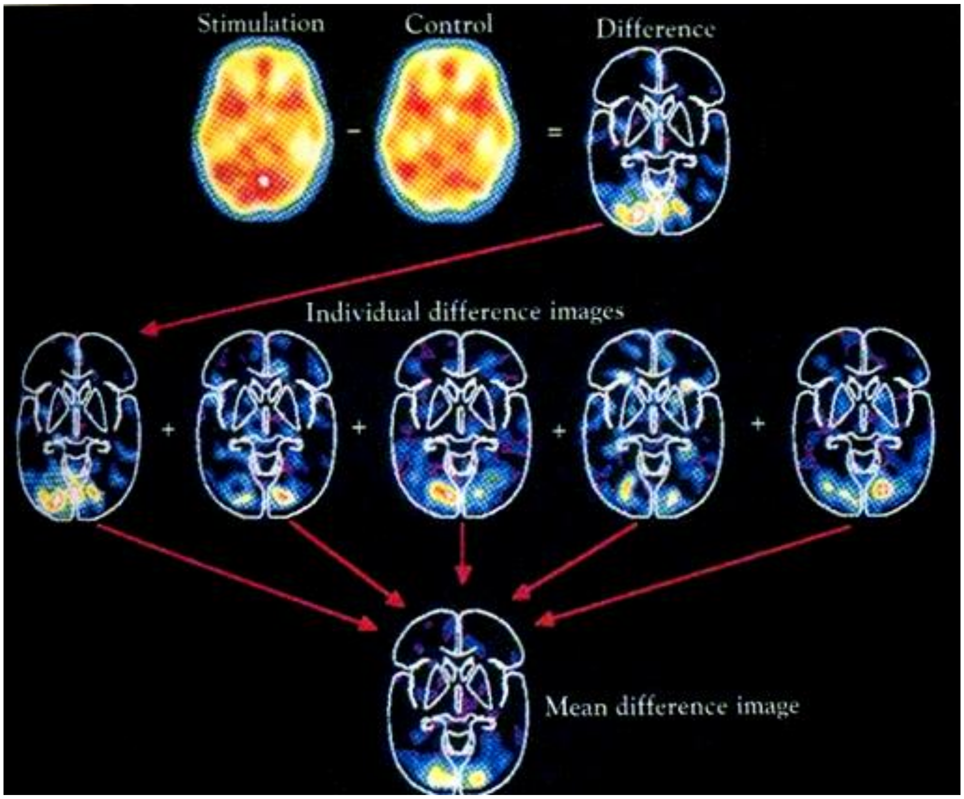

Figure 4. Representative approach to fMRI group studies
(Source: Culham, 2008, University of Western Ontario)

It can be easily observed how different the mean difference image is from the others in the group experiment. This oversimplified example of the issues with group studies serves to demonstrate the problem with making any assumptions based on the group averages. In this simple example, at least 5% of the brain is active in the mean difference image (which is only two-dimensional), while in most of the individual subjects, far less of the brain appears to be active. The primary mechanism for recovering significance from these studies is to employ sophisticated statistical software to the data. As mentioned earlier, this is an area of ongoing research.

Examples and the Ethics of fMRI

Many topics are studied with fMRI that have direct bearing on the idea of ethics. Case studies abound. Some researchers are using it to diagnose mental illness (Sorensen, 2006; Harrison et al., 2007; Kumari et al., 2006), while some are specifically measuring psychopathy in prison inmates (Harenski et al., 2010, Kiehl, 2006). Many are looking into its ability to assist with legally determining guilt or innocence, while others are investigating its use it in sentence mitigation (Tovino, 2007). Still others are detecting pain and malingering (Kolber, 2007; Grey, 2007). Some are attempting to determine the extent (or absence) of our free will (Kaposy, 2009; Greene & Cohen, 2004; Morse, 2008; Klein, 2002; Danquah et al., 2008). Some are looking more directly into human consciousness (Pockett, 2002; Pockett, 2004) while others explore the neural basis of moral cognition (Moll et al., 2005; Moll et al., 2008). Still others are measuring neural correlates for religion and belief (Harris, Sheth, & Cohen, 2008; Harris et al., 2009; Kapogiannis et al., 2009; Harris, 2010). These are just a few of the topics being studied.

It is well beyond the scope of this paper to discuss all the ethics-relevant research to which fMRI has been applied. Many have pointed out that whenever fMRI is in use, ethics questions arise (Fuchs, 2006; Farah, 2005; Rosen & Gur, 2002; Buller, 2006; Racine, Bar-Ilan & Illes, 2005; Illes & Bird, 2006; Cheung, 2009). So, instead of repeating an extensive literature and attempting to structure a broad-based discussion of ethical issues related to fMRI, I will present a few brief examples. These cover the terrain from "positive and ethical," to "needs further study," to "a real time-waster" and provide a reasonable survey of the range of issues being explored with fMRI.

fMRI: Promising for Pre-surgical Mapping

It is important to understand that even though fMRI is an art, it is a powerful one. It has been demonstrated to be used effectively in the area of pre-surgical mapping of brain regions to ensure that tumor removal can proceed without risk of damaging primary motor, visual and auditory, speech and language processing. Minimally invasive resection of brain tumors aims to remove as much of the affected tissue as possible, while preserving essential brain functions. fMRI can be used to effectively plan trajectories for surgical interventions while saving as much of the brain as possible (Vitali et al., 2011; Romano et al., 2009; Tieleman et al. 2009).

The protocol can be described simply. With the patient in the fMRI, a trained technician has them perform specific motor processing or language processing activities. Active areas near a tumor can be preserved by planning a surgical route that avoids damaging them. Importantly, as Leslie Baxter of the Barrow Neurological Institute (BNI) reminds us, fMRI can only reveal what parts of the brain are involved in specific tasks. It cannot reveal what parts of the brain are necessary or sufficient for those tasks and neither can it tell us what parts of the brain are *not* involved in a specific task (Baxter, 2011). Even the magic of fMRI cannot remove all risks from brain surgery.

Despite the many successes at BNI, brains are complicated. Haller & Bartsch (2009) are quick to warn that even pre-surgical mapping has its risks. Even if a pre-surgical mapping is entirely correct, the brain surface shifts dramatically after opening the skull and removing the cranial dura mater. Such deformation can result in up to 1 cm change in the pre-surgical mapping and despite sophisticated tools and techniques (such as a biomechanical simulation of the deformation in the navigation system) to accommodate such changes, it is often difficult to determine what portions of the brain are being impacted as surgery progresses. Note that these complexities arise for skilled surgeons and technicians that are using fMRI to "see" what is actually there on a single patient at a given point in time. Imagine the difficulty in extrapolating data across multiple subjects in a group study. Still, pre-surgical brain mapping can be very effective in planning surgical routes and there have been many reported successes in this area.

fMRI: Questionable for Lie Detection

Has science discovered "neural correlates" for lying? Can we watch a brain "light up" through the clever contrivances of software and know when deceptive thoughts are brewing? Has modern science finally deployed a more effective (albeit more costly) polygraph machine? Some say yes, and they are already marketing their solutions.

Daniel Langleben of the University of Pennsylvania has licensed his method (Davatzikos, 2005; Langleben et al., 2005; Langleben, 2002) to a company called NoLieMRI which claims "current accuracy is over 90% and is estimated to be 99% once product development is complete" (*noliemri.com*, visited April 17, 2011). Oddly, they also boast that their process is "insensitive to countermeasures by suspects" yet shortly thereafter cite a *limitation* of their process as “individuals cannot move around during the MRI scanning process” suggesting very simple countermeasures to a potential subject. Apparently, as long as a suspect wants to get caught in a lie and chooses to remain still, NoLieMRI can catch them. But, despite such inconsistencies, NoLieMRI is moving ahead with their marketing plans.

Frank Kozel of the University of Texas licensed his method (Kozel, Padgett & George, 2004; Kozel et al., 2005) to Cephos Corporation. Their website announces “Cephos has documented over 97% accuracy in blind clinical testing” (*cephoscorp.com*, visited April 17, 2011). They also boast “the only exclusively owned, issued fMRI patent in this field” though it is unclear why an “exclusively owned” patent is better than one shared with the principle investigators unless it just means individual investors can earn more profit. Nevertheless, Cephos puts on an excellent front with their beautiful website and marketing prowess.

It is clear, then, that there are firm believers in this technology. fMRI lie detection will likely soon turn into big business. But as mentioned earlier, some think we’re dabbling in a new phrenology (Uttal, 2001; Poldrack, 2010). While toeing a fine line between support for the research and involuntarily discrediting the research outright, Sean Spence of the University of Sheffield suggests that a much more careful approach must be taken.

Spence has long played foil to the likes of Langleben (Spence et al., 2004). Interestingly, within three years of the original publication of Langleben et al. (2005), two follow-up articles were specifically requested by the editors of *Legal and Criminal Psychology* in order to debate the concept of using fMRI as a lie detection scheme. Langleben was asked to support the ‘pro’ side and Spence the ‘con’ side (Langleben, 2008; Spence, 2008). In his contribution, not only were Langleben’s claims much softer than in his original 2005 research publication, but they were far more ambivalent, with success rates incomprehensibly reported as “ranging between 76 and over 90%” (Langleben, 2008, p. 4). In addition, he prefaced his remarks with “neuroscience research does not lend itself to the type of over-simplification that has plagued the interpretation of fMRI-based lie detection by the popular press and the increasingly vocal academic critics.” He also admitted to finding it a “hurdle” to “present a wholly positive view of evolving experimental data” (Langleben, 2008, p. 1). Recall that Langleben was representing the ‘pro’ side of the argument. Such warnings are appropriate when technology is this complicated. Spence’s ambivalence in being forced to take the ‘con’ side was similar, but his results were ultimately very negative. After reviewing all extant literature to date (16 peer-reviewed journal articles as of July 2007) *including two of his own papers*, Spence concludes:

> Close inspection of this literature reveals that certain central problems remain, not least the absence of replication by investigators of their own key findings. We are unable to identify a single example of this basic requirement within the extant fMRI literature (Spence, 2008, p. 24).

Note well that this warning was published by request of a legal journal, only three years ago, several years after both Cephos and NoLieMRI established their fMRI-based lie detection businesses. It appears that whether we can trust fMRI with determining the truth is a very involved question. Bets are going both ways, and the odds makers are confused.

Is fMRI ready for use in lie detection? It appears the great weight of scholarship is suggesting that it is not. But some have made strong enough cases that others have invested significant resources into making such technology commercially available. Caveat emptor.

fMRI: Useless in Determining Personhood

In an *American Journal of Bioethics* target article followed by multiple commentaries and a response, Farah and Heberlein (2007) explore the bioethical implications of scientific findings (including fMRI studies) that suggest we have an innate and genetically programmed approach to making judgments about personhood. The authors review significant findings in neuroscience and neuropsychology and discuss various definitional criteria for personhood in light of such data. They conclude that personhood is an illusion and argue that in light of this, we must adopt a strictly utilitarian approach to ethics.

After summarizing humankind's historical efforts to philosophically and psychologically define the concept of personhood and finding these age-old debates insufficient to define personhood, they turn to neuroscience to see what it brings to the debate. The bulk of the article is consumed with their introduction and discussion of neuroscientific and neuropsychological studies that demonstrate definite neural correlates of personhood and show that the concept of personhood is, in fact, innate. Oddly, they do this only after tipping their hand that they expect to be unsuccessful in defining personhood via empirical science: "We believe that this empirical, neuroscience-based approach to defining personhood will eventually be successful in translating the psychological criteria discussed earlier into neurological criteria. In so doing, however, it will be *equally successful* as the psychological approaches, *not more successful*" (Farah and Heberlein, 2007, p. 40, emphasis added). I think they put this claim up front to ensure the reader doesn't make the mistake of too hastily concluding what they will suggest is the *wrong* thing from the research findings presented.

But they go much farther. Not only do they predict neuroscience will add no value to the discussion, they presage their conclusion so we have no way of receiving the empirical data with an open mind:

> The real contribution of neuroscience to understanding personhood may be in revealing not what persons are, but rather why we have the intuition that there are persons. Perhaps this intuition does not come from our experiences with persons and non-persons in the world, and thus does not reflect the nature of the external world; perhaps it is innate and structures our experience of the world from the outset. Thus, instead of naturalizing the concept of personhood by identifying its essential characteristics in the natural world, neuroscience may show us that personhood is illusory, constructed by our brains and projected onto the world (Farah and Heberlein, 2007, p. 40).

It is unclear how the authors can conclude that "innate" does not mean "naturalizing" (especially given that the English words share the same root and are often used synonymously), but it is more unclear why they would seed the discussion thus at the outset. It is an odd line of reasoning they set forth, roughly equivalent to: It's built-in, innate, and hard-wired, so it's *not real*. The leap of logic between "perhaps it is innate" and "personhood is illusory" leaves the reader confused over how receptive or critical he should be to the forthcoming scientific analysis.

Nevertheless, Farah and Heberlein proceed to prove the innateness of the person identification reflex. With compelling evidence they demonstrate that personhood is, in fact, a built-in concept; that it's genetically determined. They first spend considerable time outlining the brain areas involved in what they call the "person network" which others refer to as "the social brain" (Farah and Heberlein, 2007, p. 42, see Figure 5). Then, they clearly depict the person network as having remarkable automaticity (triggered by external stimulus) and innateness (genetically preprogrammed and available at birth). Finally, they summarize their findings with the news that "specialization for persons comes about prior to experience with persons and other objects in the world.... prior to virtually any opportunity to learn.... In other words, a certain region of cortex is destined for face recognition as early as age 1 day" (Farah and Heberlein, 2007, p. 43).

Farah and Heberlein's science is quite interesting and illuminating. While only one commentator (Phelps) responds with a specifically scientific argument, it does force fairly extensive revision to the authors' work and draws attention to the difficulty in specifying "neural correlates" for any processing in the brain. Phelps disagrees with the specifics of what constitutes the "person network" in the brain, though, in the end, suggests that it's "just quibbling about the details" (Phelps, 2007, p. 50). Alas, details are important, so let's explore it a bit more closely.

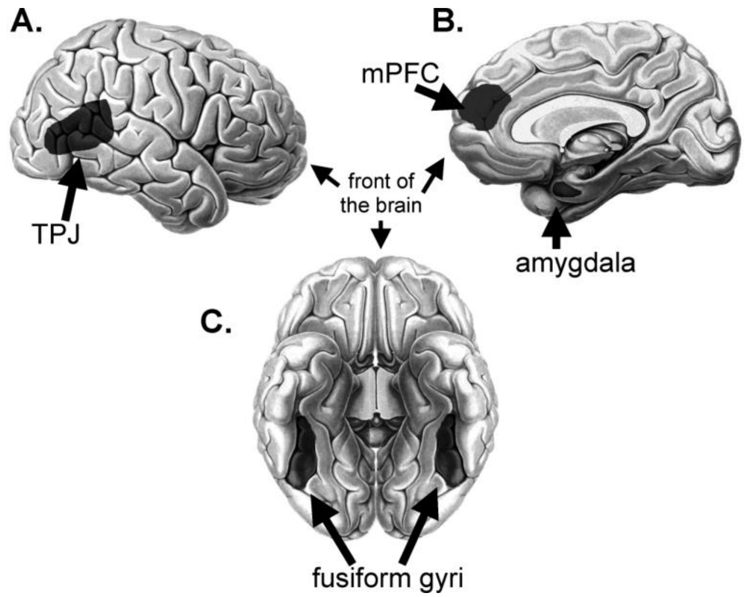

Figure 5. The Personhood Network (ala Farah & Heberlein)
(Source: Farah and Heberlein, 2007, *American Journal of Bioethics*)

Farah and Heberlein define the neural correlates of the person network as including four brain structures: the medial prefrontal cortex (mPFC), the temporoparietal junction (TPJ), the fusiform gyri, and the amygdala (see Figure 5). Phelps agrees that the evidence implicating the fusiform gyri in the person network is strong—it is after all, commonly known as the fusiform face area. She disagrees fairly robustly, however, about the inclusion of the other three structures in the network, arguing that "the mPFC is a relatively poorly understood region of the human brain" and that "our understanding of the precise function of the TPJ is even more limited than the mPFC" (Phelps, 2007, p. 50). Phelps supports her case well, citing several recent and very germane scholars. With respect to the amygdala, after a brief review of its "limited role in social responding in humans," Phelps concludes "it is hard to argue that the amygdala possesses one of the primary features of the person network" (Phelps, 2007, p. 50). One might wonder how such a significant disagreement can be anything but devastating to Farah and Heberlein's argument. The answer lies simply in assuming that if there *are* unique neural correlates to person identification in the brain, they are simply far less extensive than Farah and Heberlein claim. While their claim has less support, it is not completely unwarranted though much doubt has been cast.

Farah and Heberlein conclude from this that, since a definition of personhood has eluded us even though the concept is clearly supported in our brains, personhood must be an illusion. Despite all their scientific arguments for a person network in the brain, they conclude science "has yet to identify useful criteria for personhood" (Farah and Heberlein, 2007, p. 44). While that might be true, we must remember that they haven't really proven or disproven anything. What they *have* done is to demonstrate that the brain exhibits unique activation patterns over the concept of personhood—grounding the concept of personhood in reality. Frankly, this evidence seems to beg for the opposite conclusion.

What is most interesting is the extent of disagreement between scholars who have devoted their lives to the study of these mechanisms. It seems clear that this is still a nascent field of study which invites much scholarly response. But Nelson, for example, was unconcerned about the internal bickering of a few scientists. While he did not dispute their formulation of the person network, he did respond to the authors' conclusions drawn from the science: "Farah and Heberlein fail to make a good case for any of the major consequences they extract from the scientific results. They fail to provide us with good reason to think of persons as illusions in any ordinary sense of illusion" (Nelson, 2007, p. 65). And while Nelson was unconvinced, Perring was downright disturbed by the authors' verdict. He says "what is most problematic in the target article is the use of scientific verification to determine reality, and the corollary that if personhood cannot be determined scientifically then it is not real. In true positivistic fashion, they even suggest that philosophical debates about personhood are pointless" (Perring, 2007, p. 67). These are strong words from a philosopher who thinks the debate might have been closed too soon.

On the other hand, Sagoff seems to wave off the science entirely when he argues that "among the many senses of the term personhood, three—legal, psychological and moral—may be the most important. We can usefully and consistently project the concept in each of these senses *without any help from neurological or biological science*" (Sagoff, 2007, p. 72, emphasis added). According to Sagoff, the disagreement between the scientists was ancillary. And, though Roskies would no doubt agree (with Nelson and Sagoff) that the authors did no harm to the concept of *personhood*, she goes one step further and argues that Farah and Heberlein have *enriched* the concept of personhood with their findings: "while the things that matter may still be up for debate, continuing the debate may lead us to develop a richer notion of *personhood*, one that admits of degrees and kinds. That alone strikes me as a huge step forward" (Roskies, 2007, p. 57). Perhaps the debate isn't closed after all, instead the authors have done no harm *and* they've gotten people talking about it.

Before addressing the ethical concerns, a brief review is in order. First Farah and Heberlein made a casual pass through our history of defining the concept of personhood and concluded that "it's tough to provide a definition that outlines both necessary and sufficient criteria." The commentator's reaction was an unenthusiastic "what's new?" Second, they took a provocative spin through neuroscience, proposed neural correlates for a person network and concluded that personhood is an illusion. With this, the commentators disagreed strongly. Finally, in turning their sights toward ethics, the authors conclude "the only alternative we can see is a shift to a more utilitarian approach" (Farah and Heberlein , 2007, p. 46). Here, we will see, the commentators had equally strong negative responses.

In response to Farah and Heberlein's claim, Grey, Hall, and Carter suggest "utilitarianism can only be developed on the basis of an understanding of pleasures or preferences being satisfied (or thwarted), and these basic psychological notions face exactly the same problems of identification as the notion of a person" (Grey, Hall, and Carter, 2007, p. 58). With this they remind us that if Utilitarianism is about maximizing overall well being, then how are we expected to define this concept, or any other psychological notion for that matter? Isn't "well being" personal? Meghani carries the flag a bit further. Referring to it as a "critical flaw," she echoes the remarks of Grey, et al.: "one cannot hold that personhood, defined for purposes of this critique as the possession of certain psychological capacities, is an illusion, and in the same breath, insist that society ought to focus on fostering those very psychological traits" (Meghani, 2007, p. 53). Glannon seems to put a final nail in the coffin: "if personhood is an illusion and not an ontological category, then it is unclear how Utilitarianism, or indeed any ethical theory, could be viable and adequately ground the obligation to maximize the good" (Glannon, 2007, p. 69).

Indeed, Farah and Heberlein have placed themselves between a rock and a hard place. In declaring personhood an illusion, they seem to have destroyed a basis—or at least a primary driver—for ethics. Then, in a feeble attempt to recover ethics from an existential demise, they conclude that Utilitarianism is the only viable choice. Immediately one is forced to ask: if personhood is truly an illusion why must we recover ethics at all? Is the need for corporate morality a *stronger* illusion than the need for personhood? Farah and Heberlein's conclusion seems to be an equivocation. Clearly there is tension between what science is saying and what we believe about ourselves. It seems that the more science teaches us about our humanity, the less human we feel. Must science be dehumanizing? Racine summarizes the tension well with a veiled accusation pitched as a warning: "how will we avoid—while keeping in mind the interest of neuroscience research—disseminating forms of scientism and technological fix that reduce our take on the individual in society?" (Racine, 2007, p. 75). In effect, he's asking: how can we balance the seemingly authoritative explanatory power of science with all we hold dear as humans? How indeed?

I would like to close with some personal reflections on Farah and Heberlein's idea of "locating" personhood with an fMRI and then promptly dispatching it along with all person-based ethical frameworks. To recapitulate, Farah and Heberlein formulated two conclusions: (1) personhood is an illusion, and (2) its illusory nature undercuts ethical systems based on personhood. With respect to their first conclusion, I think "emergent" would have been a far better descriptor than "illusory." Personhood is emergent. Emergentism suggests that the mind is a complex entity. Nearly all of us could agree that personhood is an emergent feature of the human brain. That there seems to be neural correlates for personhood simply confirms that our gray matter is at least *part* of who we are. But, we are far more than our gray matter. Like any complex system, the mind exhibits behaviors and characteristics that are greater than the specific sum of its parts. Walter Glannon puts it this way:

> The mind emerges from and is shaped by interaction among the brain, body, and environment. The mind is not located in the brain but is distributed among these three entities (Glannon, 2009, p. 321, see also Glannon, 2007).

This "extended mind thesis" originated with Clark (1997), was clarified by Clark and Chalmers (1998), and was ratified by the work of Levy (2007a, 2007b). The concept was clearly available to Farah and Heberlein and should have been given some consideration. Adopting such language in their first conclusion might have also prevented Farah and Heberlein from drawing their second conclusion. In fact, it might have completely reversed the conclusion and resulted in something akin to: because the emergent properties of the brain result in our conceptualization of personhood, we understand the very human compulsion to discuss and implement a diverse array of ethical constructs. It seems that this might have been a more intellectually honest approach to the findings.

Second, and following closely on the heels of my presumptive restatement of their second conclusion, I want to draw attention to part of their neuropsychological research that did not enter the dialogue with the commentators and seems to have been given short shrift. Citing Kulhmeier et al. (2003) and Bloom (2004), Farah and Heberlein point out that research indicates an infant's behavior "reflects the child's assumption that *the important part of a person is the nonmaterial part*" (Farah and Heberlein, 2007, p. 43, emphasis added). This is not just compelling support for my argument for use of the term "emergent." It goes deeper in that it illustrates that just because something is nonmaterial does not mean it is not real. That our emergent concept of personhood also drives us to realize that the important part of a person is the nonmaterial part deserved a more prominent mention in the dialogue—and opens the door to far more discussion (which is beyond the scope of this paper).

In conclusion, Farah and Heberlein's convincing argument for the innateness of the person network (despite severe attenuation of the impact of the fMRI results ala Phelps) seems to instill a definite-ness into personhood as opposed to making it illusory—though, with Phelps (2007), we must be agreeably careful about adopting the proposed neural correlates. Since personhood is built-in, this means it is not simply memetic (Dawkins, 1976; Blackmore, 1999), though philosophical history has layered many ideas on the construct. No, personhood is more than an illusion. In fact, it's more than just an idea that has been passed down to help us deal with life (like some speak of morality or religion). It's a built-in reality we must acknowledge. And this reality lends itself to ongoing ethical discussion for centuries to come—fMRI or not.

Conclusion

In 2007, Hank Greely and Judy Illes published an article in *The American Journal of Law and Medicine* which today stands as probably still the best (and most damaging) regarding the utility of state-of-the-art fMRI. Their discussion was largely focused on courtroom use, but their analysis can be applied more broadly than the legal realm. They seem to rely heavily on Spence et al., 2004 in suggesting six primary areas of concern regarding fMRI (which I have augmented with notes pointing to discussions in *this* work):

1. A lack of individual results and a focus on "group" studies (see previous discussion of concerns related to group studies),
2. Failure to reliably repeat experiments and replicate results (see discussion above of Spence's findings on lie detection),
3. Use of small and nondiverse groups that are not representative of the general population (the largest segment of the population tested in fMRI studies is, you guessed it, young, healthy, college students who need a few bucks for the weekend),
4. Inconsistency in reporting the regions of the brain involved in mental tasking (see discussion above of neural correlates and the new phrenology),
5. Artificiality of tasks employed during testing (there is exceptionally low ecological validity when all studies are done with the subject lying perfectly still on his or her back), and
6. Lack of understanding and testing of confounding measures and other human factors (see discussion above of common misunderstandings) (Greely & Illes, 2007).

The background they provide and the depth they show in their analysis is still the best available, though now, five years hence, it is at risk of becoming dated. The recent work of Anthony Wagner of Stanford University in the *Judge's Guide to Neuroscience* reviews some later work and contributes a valuable update through 2009, but significantly, does nothing to change the overall tenor of ambivalence toward fMRI (Mansfield, Gazzaniga & Rakoff, 2010).

In summary, it behooves us to heed some simple warnings (cf. Mobbs et al., 2007) that capture the assorted issues nicely and provide some very cogent assessment of the limitations of brain imaging technology:

1. fMRI is not mind reading. Not only can it not tell us what a subject may have been thinking when a particular act was performed, it cannot tell us what the subject is thinking while being scanned.
2. fMRI provides just one window into the behavior of an individual, and as we have seen above, opinions vary as to the value of looking through that window.
3. fMRI scan interpretation is subjective, and as such, still an art form.
4. Neural correlates for anything, whether falsehood or personhood, are imperfect and the state of this science is in flux.
5. We must be careful what we "determine" neurologically, just as we must be careful what we determine genetically (e.g., criminality), especially when based on such indirect measures of neural activity.

6. Brain images (contrived as they are) carry a lot of weight and in this case a picture may be worth well more than a thousand words.

So, what about fMRI? The physics behind NMR (and hence fMRI) is sound. Magnets work. Fields can be measured. Correlation of the BOLD response to neuronal activity is pretty good. True, we need to take care concerning temporal and spatial resolution—but even that is improving regularly. Work is also being done to combine modalities for better data; mixing EEG with fMRI has become a frequent experimental configuration (e.g., Sotero & Trujillo-Barreto, 2008). Image processing software is improving regularly as well. Preliminary findings show that there must be *some* structure-to-function mapping within the brain and work is progressing on better cognitive atlases all the time (e.g., Decety & Cacioppo, 2010; Poldrack, 2010). In the meantime, we just need to be careful when we interpret what we're seeing.

But, even if fMRI flawlessly measures the brain, does that say anything? Even if we produce a perfect cognitive map, is that everything? Are we our brains? No. We are more. Much more. Popeye got it wrong.

References

Amaro, E. and Barker, G.J. (2006). Study design in fMRI: Basic principles. *Brain and Cognition*. 60: 220–232.

Arthurs, O. J. and Boniface, S. (2002). How well do we understand the neural origins of the fMRI BOLD signal? *Trends in Neurosciences*. 25: 27–31.

Banja, J. (2007). Personhood: Elusive but not illusory. *American Journal of Bioethics*. 7(1) 60–62.

Baxter, L. (Barrows Neurological Institute) (2011, January 25). *Brain Mapping*. Lecture presented to "Law 791: Neuroscience, Ethics, and the Law" at Arizona State University, Tempe, AZ.

Beck, D. M. (2010). The Appeal of the Brain in the Popular Press. *Perspectives on Psychological Science*. 5(6) 762–766.

Blackford, R. (2007). Differing vulnerabilities: The moral significance of Lockean personhood. *American Journal of Bioethics*. 7(1) 70–71.

Blackmore, S. (1999). *The Meme Machine*. Oxford: Oxford University Press.

Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human.* New York, NY: Basic Books.

Buckner, R., Schacter, D., Andrews-Hanna, J. (2008). The Brain's Default Network. *Annals of the New York Academy of Sciences*. 1124(1) 1-38.

Bufford, C. and Allhoff, F. (2007). Neuroscience and metaphysics (redux). *American Journal of Bioethics*. 7(1) 58–60.

Buller, T. (2006). What can Neuroscience contribute to ethics? *Journal of Medical Ethics*. 32: 63-64.

Brown, T. and Murphy, E. R. (2009). Through a Scanner Darkly: The use of fMRI as evidence of Mens Rea. *Journal of Law and Health*. 22: 319-341.

Brown, T. and Murphy, E. R. (2009). Through a Scanner Darkly: Functional Neuroimaging as Evidence of a Criminal Defendant's Past Mental States. *Stanford Law Review*. 62(2009-2010) 1119-1208.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Clark, A. & Chalmers, D. (1998). The Extended Mind. *Analysis*. 58: 7–19.

Cheung, E. (2009). A New Ethics of Psychiatry: Neuroethics, Neuroscience, and Technology. *Journal of Psychiatric Practice*. 15(5) 391-401.

Churchland, P. S. (2007). The necessary-and-sufficient boondoggle. *American Journal of Bioethics*. 7(1) 54–55.

Cox, D. D. and Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*. 19: 261–270.

Culham, J. (2008) . *fMRI for Newbies*. Retrieved April 22, 2011 from http://psychology.uwo.ca/fmri4newbies/Tutorials.html.

Danquah, A., Farrell, M. and O'Boyle, D. (2008). Biases in the subjective timing of perceptual events: Libet et al. (1983) revisited. *Consciousness and Cognition*. 17: 616-627.

Davatzikos, C., Ruparel, K., Fan, Y., et al. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*. 28(3) 663-8.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Decety, J. and Cacioppo, J. (2010). Frontiers in Human Neuroscience: The Golden Triangle and Beyond. *Perspectives on Psychological Science*. 5(6) 767–771.

Devlin, J. T. and Poldrack, R. A. (2007). In praise of tedious anatomy. *NeuroImage*. 37(2007) 1033-1041.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*. 70(5) 979−995.

de Zwart, J. A., Silva, A. C., van Gelderen, P., et al. (2005). Temporal dynamics of the BOLD fMRI impulse response. *NeuroImage*. 24(2005) 667– 677.

Diener, E. (2010). Neuroimaging: Voodoo, New Phrenology, or Scientific Breakthrough? *Perspectives on Psychological Science*. 5: 714-715.

Ellenberg, C. (2009). Lie detection: a changing of the guard in the quest for truth in court? *Law and Psychology Review*. 33: 139-149.

Fadiga, L. (2007). Functional magnetic resonance imaging: Measuring versus estimating. *NeuroImage*. 37(2007) 1042–1044.

Farah, M. (2005). Neuroethics: the practical and the philosophical. *Trends in Cognitive Science*. 9(1) 34–40.

Farah, M. and Heberlein, A. (2007). Personhood and Neuroscience: Naturalizing or Nihilating? *The American Journal of Bioethics*. 7(1) 37–48.

Farah, M. and Wolpe, P. (2004). Monitoring and Manipulating Brain Function: New Neuroscience Technologies and Their Ethical Implications. *Hastings Center Report*. 34(3) 35-45.

Farah, M., et al. (2004). Neurocognitive enhancement: what can we do and what should we do? *Nature Reviews Neuroscience*. 5 (5) 421-425.

Farwell, L. A. & Donchin, E. (1991). The truth will out: Interrogative polygraphy with event related brain potentials. *Psychophysiology*. 28(5) 531–547.

Farwell, L. A. & Smith, S. S. (2001). Using brain MERMER testing to detect knowledge despite efforts to conceal. *Journal of Forensic Sciences*. 46(1) 135–143.

Feigenson, N. (2006). Brain imaging and courtroom evidence: on the admissibility and persuasiveness of fMRI. *International Journal of Law in Context*. 2(3) 233–255.

Figley, C. R. and Stroman, P. W. (2011). The role(s) of astrocytes and astrocyte activity in neurometabolism, neurovascular coupling, and the production of functional neuroimaging signals. *The European Journal of Neuroscience.* 33(4) 577-588.

Fins, J. J. (2007). Border zones of consciousness: Another immigration debate? *American Journal of Bioethics*. 7(1) 51–54.

Fuchs, T. (2006). Ethical issues in neuroscience. *Current Opinion in Psychiatry*. 19: 600-607.

Ford, E. B. (2006). Lie detection: Historical, neuropsychiatric and legal dimensions. *International Journal of Law and Psychiatry*. 29(3) 159-177.

Ganis, G., Kosslyn, S. M., Stose, S., Thompson W.L. and Yurgelun-Todd, D.A. (2003). Neural correlates of different types of deception: an fMRI investigation. *Cerebral Cortex*. 13(8) 830–836.

Glannon, W. (2007). Persons, metaphysics and ethics. *American Journal of Bioethics*. 7(1) 68–69.

Glannon, W. (2007). *Bioethics and the Brain*. New York: Oxford University Press.

Glannon, W. (2009). Our Brains are not Us. *Bioethics*. 23(6) 321–329.

Gonsalves, B. D. and Cohen, N. J. (2010). Brain Imaging, Cognitive Processes, and Brain Networks. *Perspectives on Psychological Science*. 5: 744-752.

Grandin, T. (2011, March 1). *What makes us human? Visual thinking and different kinds of minds.* Institute for Humanities Research 2011 Distinguished Lecture given at ASU.

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

Greene, J. and Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London*. 359:1775-1785.

Grey, B. (2007). Neuroscience, Emotional Harm, and Emotional Distress Tort Claims. *American Journal of Bioethics*. 7(9) 65-67.

Grey, W., Hall, W. and Carter, A. (2007). Persons and personification. *American Journal of Bioethics*. 7(1) 57–58.

Haller, S. and Bartsch, A. J. (2009). Pitfalls in fMRI. *European Radiology*. 19 (2009) 2689-2706.

Harenski, C., Harenski, K., Shane, M., Kiehl, K. (2010). Aberrant Neural Processing of Moral Violations in Criminal Psychopaths. *Journal of Abnormal Psychology*. 119(4) 863–874.

Harris, S., Sheth, S. A., Cohen, M. S. (2008). Functional neuroimaging of belief, disbelief, and uncertainty. *Annals of Neurology*. 63:141–147.

Harris, S. et al. (2009). The Neural Correlates of Religious and Nonreligious Belief. *PLoS One*. 4(10) 1-9.

Harris, S. (2010). *The Moral Landscape*. New York: Free Press.

Harrison, B. J., Yücel, M., Pantelis, C. and Pujol, J. (2007). Task-induced deactivation of midline cortical regions in schizophrenia assessed with fMRI. *Schizophrenia Research*. 91(1-3) 82-86.

Illes, J. and Bird, S. J. (2006). Neuroethics: a modern context for ethics in neuroscience. *TRENDS in Neurosciences*. 29(9) 511-517.

Illes, J. and Racine, E. (2005). Imaging or imagining? A neuroethics challenge informed by genetics. *American Journal of Bioethics*. 5:5–18.

Johnson, M. H., Dziurawiec, S., Ellis, H., and Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*. 40:1–19.

Johnstone, R. A., and A. Grafen. (1993). Dishonesty and the handicap principle. *Animal Behaviour*. 46: 759-764.

Johnstone, R. A. (1994). Honest signaling, perceptual error and the evolution of "all-or-nothing" displays. *Proceedings of the Royal Society B: Biological Sciences*. 256: 169-175.

Kapogiannis, D. et al. (2009) Cognitive and neural foundations of religious belief. *Proceedings of the National Academy of Sciences, USA*. 106: 4876–4881.

Kaposy, C. (2009). Will Neuroscientific Discoveries about Free Will and Selfhood Change our Ethical Practices? *Neuroethics*. 2:51–59.

Kiehl, K. A. (2006). A cognitive neuroscience perspective on psychopathy: Evidence for paralimbic system dysfunction. *Psychiatry Research*. 142: 107-128.

Klein, S. (2002). Libet's Research on the Timing of Conscious Intention to Act: A Commentary. *Consciousness and Cognition*. 11: 273-279.

Koehler, R. C., Roman, R. J., Harder, D. R. (2009) Astrocytes and the regulation of cerebral blood flow. *Trends in Neurosciences*. 32:160-169.

Kolber, A. (2007). Pain Detection and the Privacy of Subjective Experience. *American Journal of Law and Medicine*. 33 (2007) 433-456.

Kozel, F. A., Padgett, T. M. and George, M. S. (2004). A Replication Study of the Neural Correlates of Deception. *Behavioral Neuroscience*. 118(4) 852–856.

Kozel, F. A., Johnson, K. A., Mu, Q., et al. (2005). Detecting Deception Using Functional Magnetic Resonance Imaging. *Biological Psychiatry*. 58: 605–613.

Kuhlmeier, V. A., Wynn, K., & Bloom, P. (2003). Attribution of Dispositional States by 12-month-olds. *Psychological Science*. 14:402-408.

Kumari, V., Aasen, I., Taylor, P. et al. (2006). Neural dysfunction and violence in schizophrenia: An fMRI investigation. *Schizophrenia Research*. 84(1) 144-164.

Langleben, D., Schroeder, L., Maldjian, J., et al. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *Neuroimage*. 15(3) 727-732.

Langleben, D., Loughead, J.W., Bilker, W.B., et al. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*. 26(4) 262-272.

Langleben, D. (2008). Detection of deception with fMRI: Are we there yet? *Legal and Criminological Psychology*. 13:1-9.

Levy, N. (2007a) Rethinking Neuroethics in the Light of the Extended Mind Thesis. *American Journal of Bioethics – Neuroscience*. 7(9) 3–11.

Levy, N. (2007b). *Neuroethics: Challenges for the 21st Century*. Cambridge, MA: Cambridge University Press.

Levy, N. (2008). Introducing Neuroethics. *Neuroethics*. 1: 1-8.

Logothetis, N. K., Pauls, J., Augath, M. et al. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*. 412: 150-157.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*. 453: 869-878.

Mansfield, A. S., Gazzaniga, M. S., and Rakoff, J. S. (Eds.) (2010). *A Judge's Guide to Neuroscience: A Concise Introduction*. University of California, Santa Barbara.

Matthews, P. M., Jezzard, P. (2004). Functional magnetic resonance imaging. Journal of Neurology and *Neurosurgical Psychiatry*. 75: 6-12.

Meegan, D. V. (2008). Neuroimaging techniques for memory detection: scientific, ethical, and legal issues. *The American Journal of Bioethics*. 8(1) 9–20.

Meghani, Z. (2007). Is personhood an illusion? *American Journal of Bioethics*. 7(1) 62–63.

Meyers, C. (2007). Personhood: Empirical thing or rational concept? *American Journal of Bioethics*. 7(1) 63–65.

Mobbs, D., Lau, H. C., Jones, O. D., Frith, C. D. (2007). Law, responsibility, and the brain. *PLoS Biology*. 5(4) 693-700.

Moll, J., Zahn, R., Oliveira-Souza, R., et al. (2005). The Neural Basis of Human Moral Cognition. *Nature Reviews Neuroscience*. 6: 799-808.

Moll, J., Oliveira-Souza, R. and Zahn, R. (2008). The Neural Basis of Moral Cognition: Sentiments, Concepts, and Values. *Annals of the New York Academy of Sciences*. 1124: 161–180.

Morse, S. (2008). Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience. *Minnesota Journal of Law, Science and Technology*. 9(1) 1-36.

Mumford, J. and Poldrack, R. (2007). Modeling group fMRI data. *Social Cognitive and Affective Neuroscience*. 2: 251–257.

Mumford, J. and Nichols, T. (2009). Simple group fMRI modeling and inference. *NeuroImage*. 47(4) 1469-1475.

Nelson, J. L. (2007). Illusions about persons. *American Journal of Bioethics*. 7(1) 65–66.

Norman, K. A. et al. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*. 10: 424–430.

Pellerin, L., (2005). How astrocytes feed hungry neurons. *Molecular neurobiology*. 32(1) 59-72.

Perring, C. (2007). Against scientism, for personhood. *American Journal of Bioethics*. 7(1) 67–68.

Phelps, E. A. (2007). The neuroscience of a person network. *American Journal of Bioethics*. 7(1) 49–50.

Pockett, S. (2002). On Subjective Back-Referral and How Long It Takes to Become Conscious of a Stimulus: A Reinterpretation of Libet's Data. *Consciousness and Cognition*. 11(2) 144–161.

Pockett, S. (2004). Does Consciousness Cause Behaviour? *Journal of Consciousness Studies*. 11(2) 23–40.

Poldrack, R. A. (2010). Mapping Mental Function to Brain Structure: How Can Cognitive Neuroimaging Succeed? *Perspectives on Psychological Science*. 5(6) 753-761.

Poldrack, R. A., Halchenko, Y. O. and Hanson, S. J. (2009). Decoding the Large-Scale Structure of Brain Function by Classifying Mental States Across Individuals. *Psychological Science.* 20(11) 1364-1372.

Poldrock, R.A. (2008). The role of fMRI in Cognitive Neuroscience: where do we stand? *Current Opinion in Neurobiology*. 18:223-227.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*. 10(2) 59–63.

Racine, E., Bar-Ilan, O., Illes, J. (2005). fMRI in the Public Eye. *Nature Reviews Neuroscience*. 6: 159-164.

Racine, E. (2007). Identifying challenges and conditions for the use of neuroscience in bioethics. *American Journal of Bioethics*. 7(1) 74–76.

Ramsey, J.D., Hanson, S.J., Hanson, C., et al. (2010). Six problems for causal inference from fMRI. *NeuroImage*. 49(2) 1545-1558.

Raichle, M., MacLeod, A., Snyder, A., et al. (2001). A Default Mode of Brain Function. *Proceedings of the National Academy of Sciences (USA)*. 98(2) 676-682.

Romano, A., D'Andrea, G., Minniti, G. et al. (2009). Pre-surgical planning and MR-tractography utility in brain tumour resection. *European Radiology*. 19(12) 2798–2808.

Rosen, A. C. and Gur, R. C. (2002). Ethical considerations for neuropsychologists as functional magnetic imagers. *Brain and Cognition*. 50: 469–481.

Roskies, A. L. (2007). The illusion of personhood. *American Journal of Bioethics*. 7(1) 55–57.

Roskies, A. L. (2008). Neuroimaging and Inferential Distance. *Neuroethics*. 1: 19-30.

Roskies, A. L. (2009). Brain-Mind and Structure-Function Relationships: A Methodological Response to Coltheart. *Philosophy of Science*, 76: 927–939.

Rowell, J. T., Ellner, S. P., and Kern Reeve, H. (2006). Why Animals Lie: How Dishonesty and Belief Can Coexist in a Signaling System. *The American Naturalist*. 168(6) E180-E204.

Saghoff, M. (2007). A transcendental argument for the concept of personhood in neuroscience. *American Journal of Bioethics*. 7(1) 72–73.

Seixas, D., Basto, M.A. (2008). Ethics in fMRI Studies: A Review of the EMBASE and MEDLINE Literature. *Clinical Neuroradiology*. 18:79–87.

Sofroniew, M. V. (2010). Astrocytes: biology and pathology. *Acta neuropathologica*. 119 (1) 7-35.

Sorensen, G. (2006). Future prospects for fMRI in the clinic. *Journal of Magnetic Resonance Imaging*. 23(6) 941-944.

Sotero, R. C. and Trujillo-Barreto, N. J. (2008). Biophysical model for integrating neuronal activity, EEG, fMRI and metabolism. *NeuroImage*. 39: 290–309.

Spence, S. A., Farrow, T. F. D., Herford, A. E., et al. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport*. 12(13) 2849–2853.

Spence, S. A., Hunter, M. D., Farrow, T. F. D., et al. (2004). A cognitive neurobiological account of deception: evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London B*. 359: 1755-1762.

Spence, S. (2008). Playing Devil's advocate: The case against fMRI lie detection. *Legal and Criminological Psychology*. 13: 11–25.

Stephan, K., Penny, W., Daunizeau, J., et al. (2009). Bayesian model selection for group studies. *NeuroImage*. 46: 1004–1017.

Stevens, W. D. et al. (2008). A Neural Mechanism Underlying Memory Failure in Older Adults. *The Journal of Neuroscience*. 28(48) 12820–12824.

Stoller, S. E. and Wolpe, P. R. (2007). Emerging Neurotechnologies for Lie Detection and the Fifth Amendment. *American Journal of Law and Medicine*. 33: 359-375.

Tagamets, M. A. and Horwitz, B. (2001). Interpreting PET and fMRI measures of functional neural activity: the effects of synaptic inhibition on cortical activation in human imaging studies. *Brain Research Bulletin*. 54: 267–273.

Tieleman, A., Deblaere, K., Van Roost, D., Van Damme, O., Achten, E. (2009). Preoperative fMRI in tumour surgery. *European Radiology*. 19(10) 2523–2534.

Toga, A. W. and Thompson, P. M. (2007). What is where and why it is important. *NeuroImage*. 37(2007) 1045–1049.

Tovino, S. (2007). Functional Neuroimaging and the Law: Trends and Directions for Future Scholarship. *The American Journal of Bioethics*. 7(9) 44-56.

Trovillo, P. V., (1939). A History of Lie Detection. *Journal of Criminal Law and Criminology*. 29(6) 848-881.

Uttal, W. R. (2001). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge, MA: MIT Press.

Uttal, W. R. (2002a). Précis of The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain. *Brain and Mind*. 3: 221–228.

Uttal, W. R. (2002b). Functional Brain Mapping – What Is It Good For? Plenty, but Not Everything! *Brain and Mind*. 3: 375–379.

Van Horn, J. D. and Poldrack, R. A., (2009). Functional MRI at the crossroads. *International Journal of Psychophysiology*. 73: 3–9.

Vitali, P., Dronkers, N., Pincherle, A., et al. (2011). Accuracy of pre-surgical fMRI confirmed by subsequent crossed aphasia. *Neurological Sciences*. 32(1) 175–180.

von Hippel, W. and Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*. 34: 1–56.

Vul, E., Harris, C., Winkielman, P. and Pashler, H. (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspectives on Psychological Science*. 4(3) 274-290.

Warwick, J. M. (2004). Imaging of Brain Function Using SPECT. *Metabolic Brain Disease*. 19(1/2) 113-123.

White, A. E. (2010). The Lie of fMRI: An Examination of the Ethics of a Market in Lie Detection Using Functional Magnetic Resonance Imaging. *Healthcare Ethics Committee Forum*. 22: 253–266.

Wolpe, P. R., Foster, K. R. and Langleben, D. D. (2010). Emerging Neurotechnologies for Lie-Detection: Promises and Perils. *The American Journal of Bioethics*. 10(10) 40-48.

Yarkoni, T. (2009). Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power. *Perspectives on Psychological Science*. 4(3) 294-298.