Bachelor Degree Project

UNIVERSITY OF SKÖVDE

The Science of Deception and fMRI Lie-Detection

Bachelor Degree Project in Cognitive Neuroscience
15 ECTS
Spring term 2014
Ruben Johansson
Supervisor: Oskar Macgregor
Examinator: Judith Annett

The science of deception and fMRI lie-detection

Submitted by Ruben Johansson to the University of Skövde as a final year project towards the degree of B.Sc. in the School of Bioscience. The project has been supervised by Oskar Macgregor.

2014-06-11

I hereby certify that all material in this final year project which is not my own work has been identified and that no work is included for which a degree has already been conferred on me.

Signature: ____

Abstract

Deception has long been of interest to humans, but only recently has the neuroscience of deception started. Similarly, lie-detection, as an applied aspect of the study of deception, has long been studied but only with the advent of imaging techniques and the development of the neuroscience of deception has it become possible to develop techniques based on scanning our brains. Currently, both areas suffer from methodological and philosophical problems. As an applied science fMRI lie-detection has greater issues to deal with, specifically legal and ethical issues. Despite interesting results, implicating frontal regions as the neural correlates of deception, the neuroscience of deception need better designs and more study to be able to draw any general inferences. By its nature fMRI lie-detection suffers greatly from this, and additional problems concerning privacy and legality make it seem too early to implement it in court or anywhere, as stated by many scientists. On the other hand the technology already exists and is likely to be used.

Keywords: deception; fMRI; lie-detection; methodology; neural correlates

Table of Contents

List of Abbreviations

ACC	anterior cingulate cortex
APFC	anterior prefrontal cortex
ASPD	anti-social personality disorder
BOLD	blood oxygen level-dependent
DLPFC	dorsolateral prefrontal cortex
EEG	electroencephalography
fMRI	functional magnetic resonance imaging
GKT	guilty knowledge test
IFG	inferior frontal gyrus
IPL	inferior parietal lobule
MFG	middle frontal gyrus
MRI	magnetic resonance imaging
PD	Parkinson's disease
PET	positron emission tomography
PFC	prefrontal cortex
SFG	superior frontal gyrus
STG	superior temporal gyrus
tDCS	transcranial direct-current stimulation
VLPFC	ventrolateral prefrontal cortex

Introduction

Being able to detect deception in others has always been important for humans in social interaction. One institution in which it is vital to find the truth is the court, highlighted in the criminal court's purpose to “establish beyond reasonable doubt” that someone is guilty. The problem lies then in finding out what the truth is. This is complicated by the fact that many people who are deceptive will want to keep that fact hidden, as deception is seen as something immoral (Bond & DePaulo, 2006). This is especially true for people involved in criminal activity will often want to disguise this fact so as not to be caught. The ability to detect deception is something we all possess to some degree, but we are generally quite bad at it (Bond & DePaulo, 2006). Because of this there have been many attempts throughout history to better detect lies.

Apart from trying to understand specific characteristic patterns that might be present when someone is deceitful (such as high pitched voice and difficulty maintaining eye-contact) there have been many more or less scientific techniques. An example is the medieval trials through fire or water. The suspect would need to carry hot iron or be thrown into water in a sack. If the suspect was not burned or did not float he or she was seen to be innocent (Ford, 2006). Later techniques (that have proven to have little-to-no evidence) include phrenology, handwriting analysis, hypnosis, and truth serums (Ford, 2006). Together with the advancement of technology more complex machines, computers, and analysis has become possible. The most famous, and of vital importance in a discussion on lie-detection, is the polygraph.

The first polygraph, introduced by William Moulton Marston in 1917, was quite a crude technique; consisting essentially of a blood pressure cuff (Ford, 2006). The assumption that Marston, and others following him, had was that an increase in blood pressure indicates anxiety, and therefore, during a situation such as questioning, deception. In 1932 John Larson added the three other physiologic parameters that are used in the modern polygraph: heart rate, respiratory rate, and galvanic skin response. The use of a polygraph generally involves an initial interview by the examiner, after which the subject is connected to the machine and the examiner delivers a series of questions. The assessment of the data is dependent on a certain assumption. The assumption is that autonomic responses indicate dishonesty, specifically lies. This connection is neither clear nor supported (Ford, 2006).

Since the introduction of the polygraph, several other lie-detection techniques have been developed. Three of the most prominent are thermal imaging, electroencephalography (EEG), and functional magnetic resonance imaging (fMRI). Thermal imaging, just like the polygraph, is based on assumed changes in the peripheral nervous system during deception. The thermal imaging technique uses a heat-sensitive camera to note heat changes in the subject's eyes during questioning. The assumption is that when the subject lies the eyes emit more heat than when they are truthful (Ford, 2006).

"Brain Fingerprinting" was developed by Emanuel Donchin and Lawrence Farwell and is an EEG based lie-detection technique. It utilizes the P300 wave that occur about one third of a second after that the subject recognizes something significant, and the following negative charge that is maximized in the frontal lobes. The technique combines EEG scanning and the questioning technique guilty knowledge test (GKT), which is based on provoking a response in the subject by introducing information only the subject (or the guilty in a criminal investigation) would know. By mixing basic questions ("what is the capital of Sweden called?") with less frequent specific questions ("were you present at the crime scene?"), and recording brain activity, the technique can supposedly tell which answers are truthful and which are deceptive. Although the technique has shown some promise further studies have been scarce due to Farwell's patent on the technique (Ford, 2006).

The most recently developed method is to utilize fMRI to identify the neural regions and networks associated with deception. Although fMRI has been used for some time within cognitive neuroscience it has been only during the last decade that there has been an interest in using it for understanding deception.

Deception can be defined in many ways, but it would be a mistake to describe deception only as a lie. Deception includes much more than only the fabrication of facts. Distortions, withholding information, learned falsehood (i.e. rehearsed falsehoods that become very natural to the deceiver), and so on, can all be included in the concept of deception. The current literature is not entirely coherent in its definition of deception, though most working in the field agree that deception includes many different kinds of cognitive processes (Abe, 2011; Christ, Van Essen, Watson, Brubaker, & McDermott, 2009). Many studies define deception loosely as a process through which one individual tries to convince another to accept something to be true that the deceiver knows to be false, most often to gain an advantage (Abe, 2011; Ford, 2006; Ito et al., 2011). The truth is seen as subjective and intentionality is often brought up as well: “deception of another individual is intentional negation of subjective truth” (Langleben et al., 2002, p. 727). Some have also included spontaneity as a requirement for real-life deception (Sip, Roepstorff, McGregor, & Frith, 2008), but only in the sense that real-life deception is not externally instructed or cued. In this essay deception will be handled as intentional and subjective, and will be mainly concern simple fabrications (i.e. a lie, such as saying no instead of yes to recognizing a face) as this is the most common deception in the studies presented and discussed. However, the core of the definition involves the deceiver being intentionally and subjectively untruthful (thus including manipulations that are not necessarily simple fabrications, i.e. manipulating features of a statement so that it is neither entirely truthful nor false). It should be noted that this definition is also problematic as deception can involve telling the truth while intentionally seeming to tell a lie.

The study of deception can take many forms. For example, Greene and Paxton (2009) studied the neural activity of honest and dishonest decisions, while most of the studies used in this essay deal with the act of deception rather than the decision (e.g. Abe, Suzuki, Mori, Itoh, & Fujii, 2007). Baumgartner, Fischbacher, Feierabend, Lutz, and Fehr (2009), on the other hand, investigated the neural activity correlated with breaking a promise, arguably a type of deception. Two arms of research on the neuroscience of deception have developed: one that studies specifically the neuroanatomical and cognitive aspects of deception and one that tries to refine fMRI lie-detection techniques (Langleben, 2008). Note that though these two arms of research are often linked fMRI lie-detection is inherently dependent on findings within the neuroscience of deception, but not the other way around. The neuroscience of deception is a broad research area while fMRI lie-detection is an application of the findings. Research in one of the areas need not affect the other, though it often does.

Deception in court will often be related to remembering past events. Although there might be cases in which it is clear that the person remembers the event clearly but purposefully deceives, there is a large body of evidence on the problems of eye-witness testimony (Abe et al., 2008). We humans are very good at automatically and unconsciously constructing false memories and remembering false details. This problem will not be discussed in any length in this essay, but for any attempt at clinical or scientific detection of deception this will have to be dealt with.

While fMRI lie-detection is a very new technique it should be noted that neurology is not new in court. A clear example is when magnetic resonance imaging (MRI) scans were used as supporting evidence to prohibit capital punishment of juveniles in the US because their prefrontal lobes could not be deemed to be mature enough for them to take full responsibility for their actions (Garland & Glimcher, 2006; Rusconi & Mitchener-Nissen, 2013). In India a technique based on EEG was used to convict a suspect of murder (Abe, 2011). The evidence was seen as so strong that the "judge cited the EEG results as proof that the suspect had specific knowledge about the crime that only the murderer could have and eventually imposed a life sentence" (Abe, 2011, p. 570). In addition, though currently fMRI may not be used to any large degree in courts (Rusconi & Mitchener-Nissen, 2013), the possibility that it will be used in other areas (including job interviews, pre-formal hearings, sentencing phases, and so on) is much greater and shows the need for greater understanding (Garland & Glimcher, 2006). A clear indication of that the technique is already being used is that there are two commercial companies in the US that provide fMRI lie-detection.

By combining fMRI and our understanding of the neuroscience of deception a new type of lie-detection has been developed. The technique is still very new and controversial, but it is important to study and discuss fMRI lie-detection, for reasons that will be made clear in this essay. Many limitations exist, but in this essay there will be a focus on the methodological limitations of fMRI lie-detection and on reviewing the neuroscience behind the lie-detection (not only for its importance for the technique, but because it is an interesting area in itself). It should be pointed out that by its nature the neuroscience of deception is a broad area of study, with fMRI lie-detection being a specific application of the results found within the research area. Though the neuroscience of deception can provide support for fMRI lie-detection it is not necessarily so.

Thus, the aim of this essay is to discuss the findings of the neuroscience of deception, and its application in fMRI lie-detection. To be able to determine the plausibility of fMRI lie-detection the findings of the neuroscience of deception as a whole need to be taken into account, also, non-neurological limitations of the technique need to be discussed, even though the focus will be on the methodological limitations.

A short historical background has been given in the introduction. In in the next section the neuroscience of deception will be discussed, with a main focus on the findings, as well as the connection between deception and other cognitive processes. In the section “Lie-detection Using fMRI” fMRI lie-detection will be discussed, with a focus on the fMRI technology, studies on specifically fMRI lie-detection, and methodological limitations of the technique. The following section will include a presentation of some of the limitations of fMRI lie-detection that are not methodological or technological in nature. Lastly, a discussion of the validity and reliability of the neuroscience of deception as a whole, and fMRI lie-detection specifically, will be discussed.

The Neuroscience of Deception

An often seen critique on the current studies on deception (and fMRI studies in general) is that the fact that when certain regions are activated during deception that does not mean that they are vital for deception (Rusconi & Mitchener-Nissen, 2013; Sip et al., 2008). This is underscored by the problems with ecological validity inherent in most studies on deception (Abe, 2011; Rusconi & Mitchener-Nissen, 2013). Being instructed to lie and neither having a high incentive, nor any real punishment for being deceptive is very unlike the type of deception we use in real-life situations.

One way to meet this critique is by studying patients that have lost the specific regions proposed to be implicated in deception through accidents or disease. Studying these patients provides support for theories about the neural underpinnings of deception, but also supports fMRI lie-detection as it will indicate more clearly which regions are involved or not in deceptive behavior. Up to date studies on the neuroscience of deception have used a wide variety of designs. Due to the differing designs, and often low ecological validity, it is important to be critical when reviewing the results that have been found. With this in mind it can be said that many studies have found similar results. A majority of the studies have found greater activation in deceptive conditions compared with truthful/non-deceptive conditions, with no significant neural activation when truthful/non-deceptive conditions are compared with deceptive conditions (Christ et al., 2009). Recurrent activation has been seen in the prefrontal regions, especially in regions related to cognitive control. A presentation of the current neuroscience of deception will be given in this section, the findings will be discussed more in later sections. It should be noted that though fMRI lie-detection is an application of parts of the neuroscience of deception, many (especially early studies) can be regarded as belonging in both research areas.

Early Studies

The early studies on the neuroscience of deception were often equally concerned with the neuroscience of deception as fMRI lie-detection and the assumption was that changes in the brain are involuntary (Lee et al., 2002), and theoretically we should be able to discern general anxiety from anxiety brought on by deception, and even to detect deception in people void of anxiety during deception. Spence et al. (2001) conducted the first experiment in which they used a computer-based interrogation during which participants were instructed to tell the truth or lie to specific questions. All the questions were constructed from data that the participants had given concerning everyday activities. The results of the fMRI scan showed significantly greater activation in ventrolateral prefrontal cortex (VLPFC) during deception compared to truth.

Another of the earliest studies to be done on the neuroscience of deception was a study on deception in relation to autobiographical memory and memory impairment (Lee et al., 2002). The researchers studied feigned memory impairment in six participants during fMRI scanning. The participants were shown the material and told to memorize it before the scan. During the scan they were to answer questions either truthfully or deceptively (the conditions were instructed). The deceptive condition was seen as feigned memory impairment. The deception-related activation was found in "bilateral prefrontal (BA 9/10/46), frontal (BA 6), parietal (BA 40), temporal (BA 21), and sub-cortical (caudate) regions" (Lee et al., 2002, p. 161).

Later studies have continued to find activation in these regions. Abe et al. (2008) conducted an experiment in which they posited the question whether neural activity (recorded through fMRI) can differentiate between true memory, false memory, and deception. Participants were played a list of semantically related words and later asked to judge words as old words (i.e. previously studied), semantically related non-studied words (lures for false recognition), and unrelated new words. In addition, they were to respond deceptively to half of both the old and new words. The reported activation for the main effect of deception was in bilateral dorsolateral prefrontal cortex (DLPFC), bilateral insula, right SFG, left inferior frontal gyrus (IFG), including other regions.

Similar to Lee et al., Abe et al. studied the connection between deception and memory, specifically deception and memory distortion, which both conceal the truth (but with a difference in conscious effort and intention). Abe et al. write that their most "important finding is that the left prefrontal cortex was activated during pretending to know relative to both correct rejection and false recognition" (Abe et al., 2008, p. 2817). The right hippocampus was reported to be activated significantly more during false recognition compared to denying to recognize (truthfully) and when pretending to know. Thus, their results indicated that fMRI can be used both for lie-detection as well as detecting false memories.

Langleben et al. (2005) tested the effect of attentional salience in relation to deception by using a modified version of the common questioning technique GKT. The participants were shown two cards from a normal deck of cards before being scanned and told to tell the truth about having seen one of them and lie about the other. The study had three conditions:

lie (salient), truth (salient), and repeat distracter (nonsalient truth). The repeat distracter consisted of all the other cards, which had not been pointed out before and were less salient in the situation, and these were to be answered truthfully. They found increased activation in bilateral frontal cortex, insula, inferior parietal lobule (IPL), among other regions, in the lie condition compared to the repeat distracter condition. Unlike many other studies they found no significant activation in the lie condition compared to the truth condition, even more so they found significant activation in the truth condition compared to the lie condition. In addition to contributing to the neuroscience of deception Langleben et al. (2005) aimed to test fMRI lie-detection as an applied science. They found that the regions which best predicted deception were: left IPL, left middle frontal gyrus(MFG), and left IFG. In a similar study Davatzikos et al. (2005) found deception to activate clusters centered in right IFG/superior frontal gyrus (SFG), bilateral superior temporal gyrus (STG) and IPL, and that these three clusters were the best classifiers for lie-detection (these findings will be discussed further in Lie-detection using fMRI).

Deception and Emotion

A common critique regarding the ecological validity of studies on deception is that deceptive behavior in real life is often emotional, or at least has an emotional component (Rusconi & Mitchener-Nissen, 2013). This flaw is especially significant when considering the application of fMRI as lie-detection in criminal investigations. When interrogating a suspect about an event it should be expected that both the interrogation and the possible memory of the criminal event are emotional for the suspect. Ito et al. (2011) conducted an experiment studying deceptive responses when remembering emotional or neutral pictures. They found no main effect for emotional involvement, but did find both a main effect of deception (with activations similar to other studies) as well as an interaction effect. The effect of interaction was found in bilateral DLPFC. The authors concluded that their results suggest a common prefrontal network associated with deception, regardless of the emotional valence of the stimuli. However, the ecological validity of the study can be questioned on most of the common critiques: the deceptive behavior was instructed and not spontaneous; although the pictures were emotional or neutral there was no deeper emotional involvement of the participants; the situation was artificial and not social (unlike most deceptive behavior).

Other studies have also addressed the involvement and effect of emotions on deception. Lee, Lee, Raine, and Chan (2010) studied the effect of valence of affective pictures in relation to deception. They found increased activation during both deception and truth conditions, with deception engaging a larger network during affective stimuli while truthful responses during affective stimuli produced a stronger activation in frontal and occipital regions. An important finding was that while emotional valence did affect brain activation, activity in frontal lateral, frontal medial, and parietal regions were stable during deception regardless of emotional valence (positive, negative, or neutral). Note that unlike Ito et al. (2011) the study by Lee et al. (2010) found significant activity for the main effect of emotion.

Deception and Ecological Validity

In an attempt to address four ecological limitations Spence, Kaylor-Hughes, Farrow, & Wilkinson (2008) conducted an experiment on people they knew personally. They restricted their sample to this because they included personal autobiographical data in the study. The four limitations they addressed were: spontaneity of deception (the lying was instructed, but the participants were to lie whenever they wished to, not on cue); high-cost and emotional involvement (by having their participants answer autobiographical questions that concerned events they might genuinely wish to conceal from others); vocal response[1] (rather than

[1] The possibility of confounding results due to movement artifacts is not discussed in the article. Gracco, Tremblay, & Pike (2005) outline a few designs that have been used to enable verbal report during fMRI scans. Although these methods also possess flaws Spence, Kaylor-Hughes, Farrow, et al. (2008) does not seem to acknowledge the potential design flaw.

pressing buttons, a very artificial communication); control condition (in which the participants chose whether to comply with the examiner's instructions or decline to repeat a word). They found significant activation in bilateral VLPFC in the lie compared to the truth condition, with no significant effect in reverse.

Sip et al. (2010) added aspects of social interaction, greater risk-taking, spontaneity, and a more realistic scenario for greater ecological validity. They had their participants play a dice game in which deception is an important part. The participants themselves chose when to be deceptive and when not to be. The game was a two-person game, thus allowing Sip et al. to study both the detection and production of deception, and it also added the social aspect to the study. The design included more risk-taking, not only because it was a more realistic situation (playing a game rather than responding to images being presented), but also because the game was played for money. As they did not tell the participants when to be deceptive they did not know what ratio of deceptive vs. truthful behavior there would be. The results showed that deceptive behavior was actually produced around half of the trials. The participants also accused the other player of deception about half of the time. Sip et al. only found significant activation in left premotor cortex in deceptive compared to truthful responses, but deceptive behavior compared with the control condition (in which there was no deception and participants could not see the face of the other person) showed significant activation in left prefrontal cortex (PFC), supplementary motor area, left superior parietal lobule, and the precuneus. When the truthful condition was compared with the control condition a subset of the same regions were activated, but no separate activation was found. The results are thus quite different from common activation patterns in studies on deception, but it should be noted that due to the nature of the study even the moves that were truthful included deceptive thinking (e.g. making your opponent think that you are deceptive when you in fact are not).

Sip et al. concludes that the study does not study straightforward lies or simple deception, but rather deceptive behavior in general.

PET scans allow a somewhat less artificial environment during the scanning, especially as it allows the participants to answer verbally instead of by pressing a button. In a study on deception during verbal social interaction Abe et al. (2007) found that deception (as in simple fabrication) significantly activated left DLPFC and right APFC compared to when telling the truth. The social element allowed them to see if there was a difference in brain activity if the participant was trying to deceive the interrogator (i.e. the interrogator did not know that he was being deceived) compared with simply falsifying the truth (i.e. the interrogator was aware of that the participant would be deceptive). This social element of the study showed that different regions were activated in the process of deceiving the interrogator compared with simply falsifying a truth. The process of deception (as in deceiving the interrogator) showed increased activation of left ventromedial PFC and amygdala. This type of experiment would have been hard to do using fMRI (as it used verbal communication) and shows that the type of deception used during most fMRI studies are probably not very similar to real life. A good indication of this is that Abe et al. (2007) found activation in ventromedial PFC during the process of deceiving someone, a region that has been correlated to moral behavior (Thomas, Croft, & Tranel, 2011), and deception is often seen as morally wrong (Bond & DePaulo, 2006).

Deception is a type of behavior that is often seen as morally questionable, especially in social contexts (Bond & DePaulo, 2006). Abe et al. (2007) found activation of the amygdala during deception, something Baumgartner et al. (2009) also found. Studying the neural activation during the breaking of a promise, Baumgartner et al. wished to have a design as ecologically valid as possible. Thus there was a strong social element, interaction between two real humans: the participants could choose whether to break or keep their promise. The study had three stages: in the first the participant promised another player to reciprocate, in the second the participant was informed if the other player would accept the promise and in the third they decided if they were to keep the promise or not. Honest and dishonest participants showed different neural activation in all three stages (with dishonesty generating greater activation). The promise stage showed increased activity in ACC and bilateral IFG/anterior insula in deceptive participants. The second, anticipation stage, revealed activity in right IFG and insula, again no significant activation in honest participants. The last stage had two phases: in phase A the other player's trust decision was shown, in phase B the participant was reminded of his promise. After these phases the participant either kept or broke his promise. In both phases, activation could be seen in the ventral striatum, phase A further revealed increased activity in ACC and left DLPFC, while only phase B showed activity in left amygdala. Baumgartner et al. (2009) speculate that the high ecological validity of their own study and Abe et al. (2007) is what generates the activity in the amygdala, which was expected. Thus, it is not directly deception-related activity, but rather the confrontation with another human being (Baumgartner et al., 2009) that triggered the amygdala activity.

Activation of DLPFC is common during deception and one way to test the importance of the region is through stimulation studies. Using different kinds of techniques to find the neural underpinnings of deception is likely to generate more valid results, as each technique has its own strengths and weaknesses. Mameli et al. (2010) conducted an experiment using transcranial direct-current stimulation (tDCS) targeting specifically DLPFC, to investigate if the region influences cognitive processing of deception using either general knowledge or personal information. They found that response time was greater in lies compared to truth, and lies using personal information were significantly faster than lies using general knowledge. Stimulation with tDCS sped up lies about general knowledge, but not lies about personal information and neither stimulation nor having the participant believe he received stimulation affected the response time of truth. The results of this study are unlike previous stimulation studies which have found that stimulation of DLPFC slowed lie production (Mameli et al., 2010). Stimulation of DLPFC has been shown to increase response time as in Mameli et al. (2010), but other studies have failed to give the same results (Luber, Fisher, Appelbaum, Ploesser, & Lisanby, 2009).

As the study of deception, especially in relation to lie-detection, is of special interest in relation to law some experiments study the neural correlates of deception during different situations common during criminal investigations. Bhatt et al. (2009) wished to see if they could find a different pattern of activation in deception, compared to truth, during a line-up. Specifically, they wanted to study the effect of familiarity on deception. Before the line-up the participants were shown pictures of faces. In some trials one of the faces presented was familiar (i.e. shown before the line-up), and in other trials none were familiar. Thus two different deceptions were used during the mock line-up: implicating someone unfamiliar and not implicating someone familiar. The main effect of deception showed increased activation in right frontal regions (including ACC, superior, inferior, and middle frontal gyri). Deception in relation to familiar faces activated right SFG, right IFG/VLPFC and bilateral precuneus.

Neural Substrates of Deception and Other Cognitive Processes

The regions often reported to have significant activation during deception (compared to control or non-deceptive conditions) have also been studied in relation to executive control. Christ et al. (2009) writes that there are three aspects of executive control that may contribute to deception. Working memory is important for keeping the truth in mind while formulating a deceptive response (so as to remember what is true and what is not). Suppression of a truthful response requires inhibitory control, and switching between truthful and deceptive responses requires sufficiently developed task switching. All of these three aspects have been reported to activate prefrontal regions (Abe, 2011). It can be seen from this that it is hard to differ between regions vital to deception and those not, as different types of deception, in different situations, will require different cognitive processes: deception is indeed multifaceted.

Connected to the cognitive requirements of deception Vartanian et al. (2013) conducted an experiment on the effect of workload on deception. As working memory is necessary for the production of deceptive behavior, especially coherent lies, they hypothesized that an increase in workload on working memory would decrease the participants' ability to successfully lie. The participants were instructed to memorize a series of digits which they were then to respond truthfully or deceptively to when presented with a similar or the same series in the fMRI scanner. The high workload condition was the memorization of six digits while the low workload condition included only four. The behavioral results reported were that there was a significantly higher response time in both high workload compared to low workload and lie condition compared to truth, but no interaction was reported. This is consistent with what has been found in most other studies (Abe et al., 2008; Davatzikos et al., 2005; Ganis, Rosenfeld, Meixner, Kievit, & Schendan, 2011; Greene & Paxton, 2009; Ito et al,. 2011): deceptive behavior (in this case and most other: simple fabrication of truth) elicit longer response times compared to being truthful. The fMRI scanning found an interaction between the high workload and the lying conditions, with significant activation in the right IFG. The lie condition revealed significant activation in DLPFC, SFG, the precuneus, and middle temporal gyrus. Again, these activations are similar to those reported in other studies.

In a meta-analysis of twenty-four studies on deception Christ et al. (2009) found 13 regions that were consistently activated across studies: bilateral insula, right IFG , right MFG, right IPL/supramarginal gyrus, right internal capsule/thalamus, left IFG, left IPL, left internal capsule, left precentral gyrus/MFG, right anterior cingulate, right IPL.[2] Thus, eight of 13 regions were located in or near the PFC, indicating a major involvement of the PFC in

[2] These regions include bilateral aspects of the VLPFC, DLPFC, anterior insula, and right ACC.

deception. The close connection between deception and executive control can also be seen, as 10 of the 13 regions have also been correlated with executive control. These regions have also been related to selective attention and detection of salient low-frequency target stimuli (Christ et al., 2009; see also Badre & Wagner, 2007; Tsuchida & Fellows, 2013). Christ et al. (2009) further proposes that the IPL is not related to deception, but rather detection of salience, thus exemplifying the problem of not being able to determine if a region is directly connected to a cognitive process.

In a more recent review Abe (2011) state that the most commonly reported regions are the DLPFC, VLPFC, anterior PFC (APFC) and ACC. The DLPFC has been found to be activated during monitoring and manipulation of working memory, response selection, and cognitive control (Abe, 2011). The VLPFC has been implicated in a wide range of cognitive processes, among them task switching, reversal learning, semantic retrieval, and response inhibition (Abe, 2011; Badre & Wagner, 2007). The APFC is implicated in processes associated with integration of outcomes of several separate cognitive operations in pursuit of a higher behavioral goal (Abe, 2011). ACC has also been activated during many kinds of processes, including conflict detection and emotional processing (Abe, 2011). As deception is a complicated act associated with a multitude of cognitive processes it is likely that these regions interact with each other in different forms to produce the intended deception (Abe, 2011). In addition to these regions, studies that have tried to produce more ecologically valid situations has found activation in the limbic system (Abe et al., 2007; Baumgartner et al., 2009; Rusconi & Mitchener-Nissen, 2013) but it is not clear that emotional involvement has any significant effect on neural activation during deception (Ito et al., 2011).

VLPFC has been implicated in many studies on deception and encompasses the IFG anterior to pre-motor cortex and posterior to the frontal pole (Badre & Wagner, 2007). The region, especially left VLPFC, has been found to be closely correlated with semantic retrieval as well as episodic memory retrieval. Activation of VLPFC during memory retrieval is often coupled with activation in other regions related to memory, such as lateral temporal cortex. In addition to retrieval VLPFC, especially left mid-VLPFC, has been found to be related to post-retrieval selection, which can be assumed to be very important for deception regarding past events. In relation to its' function for memory retrieval, VLPFC has been correlated with rule-guided behavior. Badre and Wagner argue that "VLPFC may be more broadly involved in retrieval and selection of representations that help to guide and constrain action through stored knowledge" (Badre & Wagner, 2007, p. 2897). One could assume that VLPFC activation during deception can be expected due to this, as a "do not lie" rule is common in most situations in society. On the other hand, the activation could also be due to compliance to experiment instructions (i.e. remembering and following previous instructions, especially during more complex designs).

A majority of the regions implicated during deception are located in the frontal cortex (Christ et al., 2009) and in comparison to studies on specifically the neuroscience of deception there is a more mature research area on the study of the cognitive functions of the frontal cortex. The prefrontal cortex has been found to activate during many different cognitive processes and has shown both general and process-specific activation. Four regions have been proposed for the former, general activation: VLPFC; DLPFC; posterior supplementary motor area; dorsal ACC (Tsuchida & Fellows, 2013). Tsuchida and Fellows investigated the effect of focal lesions on cognitive processes and found left VLPFC to be correlated with attentional shifting and the Stroop effect, left IFG with the Stroop effect, and anterior insula with attentional shifting. A similar incongruity as in the Stroop effect can be found in deception and it can thus be interpreted that the left VLPFC and IFG are important in the suppression of a truthful response and thus the production of deception. The study indicates that there is both a functional specialization as well as general processing in the PFC. This fact can be important to have in mind for fMRI lie-detection, as it is additional evidence for that there will be cognitive processes during deception that may be irrelevant.

Atypical Neurology and Deception

To validate the findings of other studies on the neuroscience of deception, but even more important for enabling valid lie-detection, the study of deception in people with atypical behavior and neurology is vital. In relation to real-life application of fMRI lie-detection the two diagnoses (and personality traits) that most need to be studied is anti-social personality disorder (ASPD) and psychopathy (and anti-social and psychopathic personality traits). Individuals with these diagnoses, or personality traits, are overrepresented in prison (Hughes, 2010; Jiang et al., 2013) and are often proficient liars (Fullam, McKie, & Dolan, 2009; Jiang et al., 2013). Fullam et al. (2009) investigated the relationship between deception and psychopathic personality trait through the Psychopathic Personality Inventory. In the fMRI study the participants were told to lie or tell the truth about simple questions (i.e. “have you made your bed today?”). The main effect of deception was found to activate VLPFC bilaterally, while truth significantly activated bilateral frontopolar PFC and bilateral medial superior frontal cortices. The authors found that Psychopathic Personality Inventory subscales had a negative correlation on activation in specific brain regions during deception: Fearlessness – right orbitofrontal; Coldheartedness – bipolar temporal lobe; Machiavellian ego – bilateral caudate; Social potency – right posterior cingulate; Stress immunity – bilateral insula. After correcting for age the negative correlation between the sub-scale and activation remained in coldheartedness and stress immunity, but not the other three. While the results of the study indicate a correlation between lower significant activation in brain regions that have been implicated during deception and psychopathic personality traits the study is still done on only college students, a possible confound.

47% of male prisoners are diagnosed with ASPD, whom, unlike typical college students, can be assumed to be very skilled at deception (Jiang et al., 2013). Jiang et al. investigated if the results from previous studies were applicable to offenders with clinical ASPD. Participants in this study were all diagnosed with clinical ASPD, but had no history of other serious mental disorders. They were divided into three different levels of deceptive proficiency (i.e. how good they were at lying). The participants were told to pick out three of ten pictures, and then instructed to tell the truth, lie (inverse of the truth), or lie with a strategy when pictures were presented to them during the scan. Deception was found to be correlated with activation in bilateral DLPFC (including MFG), left IPL (including supramarginal gyrus), and bilateral ACC/medial SFG. Thus, the offenders with clinical ASPD showed similar activation during deception compared with healthy participants in other studies. However, Jiang et al. (2013) also found that the activation of the above regions decreased as capacity for deception increased, finding a similar negative correlation as Fullam et al. (2009).

Unlike people with ASPD or psychopathy parkinsonian patients are often described as honest (Abe et al., 2009). Though it is a possibility that honest people are more likely to suffer from Parkinson’s disease (PD), the neurodegenerative effects of the illness may be what are causing the honesty. It is then a possibility that the PD patients do not choose to be honest, but rather that they have a difficulty telling lies. Using PET Abe et al. (2009) studied the difference between participants with PD and healthy participants on an instructed lie task. All the participants were instructed to verbally lie to one out of four actors, while telling the truth to the other three. The PD participants were found to be significantly worse at lying, compared to controls (i.e. they had less correct responses during the deception condition). All but one of the PD participants reported having no difficulty knowing whom to deceive, and the error in deception compliance was found to be 91% due to telling the truth (0.9% no response, 7.3% answered with both truth and lie). The PET scan showed a significant negative correlation between deception task performance and regional metabolism in right APFC and left DLPFC. Thus lower activity in DLPFC has been found to correlate both with higher ability to deceive (Jiang et al., 2013) as well as lower ability to deceive (Abe et al., 2009).

In 2007 Yang et al. conducted a MRI study in which they studied structural brain differences between three groups: controls, people with ASPD (but not rated as pathological liars), and a group consisting of people with a history of repeated lying (pathological liars). In accordance with results gained from fMRI studies Yang et al. found increased white matter in certain regions in pathological liars when compared to both control and ASPD participants. The main findings were 22-26% increased white matter in orbitofrontal cortex, 32-36% in IFG, and 28-32% in DLPFC.

No matter how sturdy results an experiment produces the study of healthy participants has the negative effect of having the possibility that the brain activation found is actually not due to what is being studied, but because of some other cognitive process. One way that clearly indicates if a region is responsible for the production of a specific cognitive process is to find out if the loss of that region hampers or causes a failure of that cognitive process to function. This can be achieved to some extent by stimulation techniques such as tDCS, but especially studying people who have suffered lesions of the specific brain region. For example, Lagemann et al. (2012) reports on a 19 year old man whose main symptom was compulsive lying and cheating. MRI scans of the man's brain revealed a complete lack of ACC, a region implicated in risk evaluation, inhibition, and regulation of behavior in relation to learned rules. Lacking the neurophysiological substrate for these processes, the man was a compulsive liar. This finding supports that the ACC is a key area in modulating deception.

In another interesting case Sellal, Chevalier, and Collard (1993) reports a case of a patient suffering from reflex epilepsy, a form of epilepsy in which a specific type of behavior provokes the attack. The 51 year old patient had 1/3 of his attacks when he lied for business reasons (the other 2/3 were due to unknown reasons). An MRI scan revealed "a meningioma, 30 mm in diameter, located on the right cavemous sinus wall and the anterior clinoid process, near the sella turcica" (Sellal et al., 1993, p. 936). The tumor compressed the right medial temporal lobe, including the amygdala. This raised the possibility that the patient felt a particular emotion when he lied and the limbic stimulation triggered the epileptic discharges. Thus, the authors called his peculiar form of reflex epilepsy "Pinocchio syndrome" (Sellal et al., 1993, p. 936).

In one of the few cases when fMRI has been used in relation to court proceedings Spence, Kaylor-Hughes, Brook, et al. (2008) studied and performed fMRI lie-detection on a woman convicted for having poisoned a child in her care. Post-verdict she had been diagnosed with Munchausen's syndrome by proxy, a syndrome with deception at the very core. The woman wished to use fMRI lie-detection to support her claim that she was innocent of the crime. Thus Spence, Kaylor-Hughes, Brook, et al. scanned her during two conditions: one in which she endorsed her accusers' version of the events (i.e. lying, if she was innocent); the other in which she endorsed her own version (i.e. telling the truth, if she was innocent). The neuroimaging data revealed significant activation in bilateral VLPFC and ACC during the deception while the truth condition only showed a small activation in left lingual gyrus. If previous studies on the neuroscience of deception are correct in implicating frontal regions in deception (especially VLPFC and ACC), these results indicate that the woman was telling the truth when she said she was innocent. The greater problem arises when discussing possible confounds: the responses may have been learned, the deep emotional involvement may not allow for previous results to be applicable to this case, and she may have used countermeasures. In addition, deceptive behavior is by definition part of Munchausen's syndrome by proxy, and considering the results found by Fullam et al. (2009) and Jiang et al. (2013) the woman may have had uncommon patterns of activation. These problems are at the very core of what needs to be discussed about fMRI lie-detection and will be the focus of the next section.

Lie-detection Using fMRI

The fMRI Technology

To be able to assess if fMRI can be a good tool for lie-detection one has to consider some of the technological and statistical basis that enables fMRI to produce real-time imaging of activity in the brain. The fMRI is an extended technique of the MRI brain scanner that allows imaging of not the structural landscape of the brain as in a standard MRI, but instead to detect changes in blood flow in the brain. The signal received, blood oxygen level-dependent (BOLD), is an indication of where there is increased activity in the brain as increased activity requires a greater supply of oxygenated blood (Langleben, 2008). During the scan the participant is placed with the head or entire body inside of an MRI scanner which generates a strong magnetic field. The scanner is equipped with a magnet that generates two types of magnetic fields, one perpendicular to the participant and the other, weaker gradient fields, at an angle. The gradient fields are rapidly switched on and off which causes the hydrogen nuclei in the body water to resonate. The resonance emits radiofrequency signals which can be detected and allows computer analysis. The analysis results in a three-dimensional reconstruction of the relative concentrations of hydrogen nuclei in the tissue. The electrical activity that happens in the brain is quick, measured in milliseconds. On the other hand blood flow is relatively slow, measured in seconds. Although this is a methodological problem of the fMRI (Farah, 2014), the coupling between neuronal activity, metabolism, and blood flow is robust (Langleben, 2008). A typical fMRI has a spatial resolution of 1-3 mm and a temporal resolution of 2-3 seconds. Both resolutions can however be modified to fit what is required (e.g. to have lower spatial resolution but higher temporal). There is no normative BOLD base-value, thus fMRI studies typically rely on the difference between a baseline control stimuli and the target stimuli. Due to this it is never exactly the same conditions being compared both within and between studies. In the context of lie-detection, one study might compare rest with deception and another deceptive response with a truthful response. This fact also makes comparison between individuals problematic as different people will not have the same baseline even if they are being measured in the same way (Rusconi & Mitchener-Nissen, 2013).

The use of fMRI has grown increasingly popular, but the technique has also been the target of much critique. Farah (2014) identifies four main groups of critique: objects of imaging; metatheoretical assumptions and goals; statistics; and influence. While a discussion of the reliability and validity of fMRI in general is beyond the scope of this essay it is important to note some of the critique, especially the “overly convincing” (Farah, 2014, p. S27) brain scans. As already stated, what is detected in an fMRI is blood flow, not brain activity. Farah admits there is a fundamental problem here, as the relationship of BOLD and neural activity is not well understood (Langleben, 2008, does not entirely agree on this), but also states that it is the best technique at hand. Some studies, especially early fMRI studies, have focused specifically on the localization of function in the brain. Due to this the use of fMRI is sometimes “caricatured as a form of phrenology” (Farah, 2014, p. S21) and although it is true that some studies focus solely on localization the parallel is probably unwarranted. Even basic research on localization is often used for the validation of methods (in the case of deception even studies focusing only on localization can sometimes contribute to better lie-detection). By its nature the use of fMRI requires extensive statistical analysis. This has sometimes been interpreted as that the images produced from fMRI is not reality, but rather a mass of statistics that do not have any real meaning. As pointed out by Farah, all science suffers from this, though fMRI may be especially reliant on statistics to represent brain activity. As fMRI is so dependent on extensive statistics there are greater problems when the statistics are incorrectly implemented. The problems of multiple comparisons and circularity have recently been pointed out by Bennett, Baird, Miller, and Wolford (2010) and Vul, Harris, Winkielman, and Pashler (2009), respectively. Farah concludes that the use of statistics is not a problem if it is done right. Especially relevant for if fMRI evidence is to be used in court is the worry that images of brain activity have too much influence: that fMRI scans are overly convincing and overly appealing. There have been studies on the influence of fMRI on juror decision-making (McCabe, Castel, & Rhodes, 2011) that found fMRI lie-detection to be significantly more appealing than evidence from thermal imaging and polygraph, but only if the validity was not questioned.[3] As pointed out by Farah, this is thus rather a question of scientific literacy in the general population, especially as there has been a failure to replicate the results from the study (Farah, 2014).

Studies on fMRI Lie-detection

One of the main arguments for implementing lie-detection technologies in court is that our ability to detect lies without help is not satisfactory. In fact our ability to detect lies has been found to be no greater than chance (Bond & DePaulo, 2006) while our ability to produce lies develops with age and use (Rusconi & Mitchener-Nissen, 2013). It can thus be argued that if we can develop reliable and valid lie-detection techniques that are significantly better than chance we ought to do so, and use them. A number of studies have studied not only the neuroscience of deception, but also its application: fMRI lie-detection. Langleben et al. (2005) as well as Davatzikos et al. (2005) used a variant of the GKT and had the participants, instructed by one experimenter, lie to the other experimenter. Both studies found similar regions to be the best classifiers for deception (though not the same regions): Langleben et al. used left IPL, left MFG, and left IFG; Davatzikos et al. used right I/SFG, bilateral STG, and IPL. Langleben et al. (2005) achieved an accuracy of 76.5%[4] (68.8% sensitivity, 83.7%

---

[3] It can be assumed that the evidence will be questioned, see Hughes (2010).

specificity) while Davatzikos et al. (2005) found their accuracy to be between 87.9-95.5% accuracy (90-90.9% sensitivity, 85.8-86.4% specificity).[5]

Participants in Kozel et al.'s (2005) experiment were similarly instructed by one experimenter to try to deceive the second experimenter in a design in which the participant was told to steal either a clock or a watch. They used three clusters for the classification, centered in: right ACC; right orbitofrontal cortex and IFG; and right MFG. Using only the right ACC cluster the accuracy was 83% (correctly identifying which item was stolen). When all three were used the accuracy rose to 90%. As part of a second experiment (Kozel et al., 2009) they replicated the ring-watch experiment from 2005 and found the accuracy to be only 71-72% (although this time the study was conducted after the first part of the experiment, thus the participants were assumed to be more fatigued). In the first part of their 2009 experiment they wished to emulate a real-life crime. To do this they included personal requests from one experimenter (seemingly apart of the study), destruction of evidence from a seemingly real crime, as well as unexpected events, and a chance of being found out. In this part of the experiment they acquired a sensitivity of 91% and a specificity of 42% when using the whole group. When using more stringent conditions for inclusion in the analysis the sensitivity rose to 100% but the specificity was only 33%.[6] The more stringent conditions were: the participants were correctly identified in the ring-watch test; and the participants correctly completed all of the protocols, had no movement artifacts, and had been determined to be either lying or telling the truth by both experimenters who were responsible for the call.

---

4 Sensitivity is the percentage of correctly determined positives and specificity the percentage of correctly determined negatives.

5 The different accuracies found by Davatzikos et al. (2005) were taken at different stages of the analysis.

6 It may be worth noting that both Kozel et al. (2005) and Kozel et al. (2009) were funded by Cephos, one of the companies offering fMRI lie-detection services.

One of the early assumed strengths of fMRI lie-detection was that since it only detects responses from the brain, someone being scanned will not be able to deceive the technique: fMRI will be resistant to countermeasures (Lee et al., 2002). This was one of the main selling points from the beginning, especially for proponents of the polygraph, as the polygraph suffered greatly from how easy it was to apply countermeasures to it. Ganis et al. (2011) studied whether fMRI lie-detection is truly resistant to countermeasures. They instructed the participants to be deceptive in certain conditions, and then instructed them on how to apply countermeasures. The countermeasure used was to move their toes and fingers (without being detected). Before the participants were instructed to use countermeasures Ganis et al. report their accuracy to have been 100%, but this changed drastically when the participants applied countermeasures; then the accuracy was only 33%.

Methodological Limitations

Rusconi & Mitchener-Nissen (2013) discusses the prospects for future use of fMRI as lie-detection. They are critical about the possibility to use the current fMRI lie-detection technology in any real-life application. In the article they bring up eight points that scientists developing the technology need to deal with. By the nature of their critique it is mainly targeted at the applied science of fMRI lie-detection, but several points are also important to consider for the neuroscience of deception in general.

Their first point is concerned with that “assumptions and inferences underlying fMRI processes and technologies need to be confirmed (or dispelled) so as to give credence to the scientific claims being made” (Rusconi & Mitchener-Nissen, 2013, p. 5). Some of these are specific to the study of deception, such as if deception is truly more cognitively demanding than telling the truth (as is assumed by many researchers), while the problems concerning reverse inferences are more general for fMRI (for a discussion see Farah, 2014). Additionally, most data analysis are dependent on knowing beforehand what is objectively defined as a lie and what is not. In forensic practice this will not be possible

The second point states that “to achieve internal validity, it needs to be conclusively determined that what is being measured is actually evidence of deception and not unrelated cognitive processes” (Rusconi & Mitchener-Nissen, 2013, p. 5). Lie-detection using fMRI depend upon the ability to detect the suppression of competing responses, but can neither determine what these are nor what the suppression implies. When being questioned it is reasonable that a suspect will think about what he says and stop himself from saying certain things, regardless of his innocence or guilt. It is also natural that the setting will evoke emotions such as fear and anxiety which will cause neural activation unrelated to deception. Unless the pattern detected can be robust for each and every individual the expert counsel will open itself up for legal challenge (e.g. see Hughes, 2010).

Third, “the question of individual differences affecting fMRI results needs to be answered” (Rusconi & Mitchener-Nissen, 2013, p. 5). It needs to be established that deception can be detected in individuals, taking into account that different brains may not process lies similarly. This includes being able to correct for variation based on age, substance abuse, mental disorders, and very high incentive to lie. Most studies have been conducted on college students (Jiang et al., 2013) and low incentive is a recurrent ecological problem (Sip et al., 2008). There have been studies on people with mental disorders (Abe et al., 2009; Jiang et al., 2013) indicating that patterns of activation during deception is similar, but the findings need to be replicated and further studied. Neurological and cognitive changes in addicts have not been studied but can reasonably be assumed to affect patterns of activation.

Fourth, “the question of whether subjects or questioners can manipulate the fMRI baseline or response data needs to be addressed” (Rusconi & Mitchener-Nissen, 2013, p. 6). In early studies (Lee et al., 2002) one of the proposed strengths of fMRI lie-detection was that it would be more resistant to countermeasures, as one was assumed to be unable to control one's cortical activity. Ganis et al. (2011) showed that countermeasures are in fact very effective. Additionally, as an interrogation with fMRI will include social interaction both experimenter and participant bias will need to be handled.

Fifth, “the subjectivity inherent in fMRI analysis algorithms needs to be acknowledged and these algorithms opened up for scrutiny” (Rusconi & Mitchener-Nissen, 2013, p. 6). There is (and will be) a major problem in that there exist several algorithms which are being used for fMRI lie-detection. As algorithms include all the subjectivity and bias of the programmers different algorithms may give different patterns of activation from the same scan. This is problematic from a scientific perspective, but disastrous if fMRI is to be used in forensic practice. If the prosecution and defense can produce different results from the same scan the probative value of fMRI lie-detection will be negated.

In the sixth point they state that “we need to determine the percentage of the population who for various reasons are unable to undertake an fMRI, as well as the nature of those reasons” (Rusconi & Mitchener-Nissen, 2013, p. 6). It would be problematic to scan people who are unable to remain still for a long time (even if they are cooperative, which cannot be assumed), as fMRI is highly sensitive to movement artifacts. Additionally, people with certain medical conditions will be unable to go into the powerful magnetic field. If too high a percentage of the population will be barred from using fMRI lie-detection its value in court can be assumed to be quite low.

The seventh point is concerned with that “questions over the methodological validity of past and future fMRI studies must be answered” (Rusconi & Mitchener-Nissen, 2013, p. 6). Few studies have reported accuracy rates and those that do differ greatly between each other in both design and results. According to Rusconi and Mitchener-Nissen (2013) the only remedy is more peer reviewed studies. In addition, many fMRI studies suffer from defective research methods (Bennet et al., 2010; Vul et al., 2009; see also Farah, 2014).

The last, eighth, point states that "to attain external validity, experiments need to be applicable beyond highly controlled laboratory settings" (Rusconi & Mitchener-Nissen, 2013, p. 6). In criminal investigations there will be very high emotional and personal stakes for those involved and the situation will be very confrontational. This is very unlike the type of situations used in current studies.

These points are in line with critique of fMRI lie-detection others have presented. Sip et al. (2008) compress their view on the problems of fMRI lie-detection into four different problems: making reverse inferences on physiology to mental states; making inferences about individuals on the basis of group studies; studying choice in the laboratory; interpretation must take into account the context. The problem of reverse inference is similar to Rusconi & Mitchener-Nissen's (2013) first and second points by focusing on that the brain activity seen during a scan during instructed deception does not by necessity need to have to do with deception. Both articles stress the importance of ecological validity, especially if fMRI is to be introduced in forensic practice.

Evidently there are several technological hurdles that must be discussed and solved if fMRI is to be used not only as a scientific tool but applied as a method of lie-detection in court. For fMRI to be a good technique to use in court it is a prerequisite that it can be applied, if not on everybody, at least on most people. However, the fact is that many people cannot be scanned in a MRI scanner. The powerful magnetic fields that are generated during the scan prohibits the presence of any metallic parts, including medical equipment, pacemakers, piercings, and so on. The nature of the machine also makes it very sensitive to movement artifacts. This means that any movement in the scanner degrades the quality of the scan (Langleben, 2008). Without discussing voluntary movement (which would be a type of countermeasure) the person being scanned will need to be still for quite a long time which in some cases can be near impossible. For example, the possibility for a person with attention-deficit/hyperactivity disorder to lie still for several minutes cannot be expected to be very high and people suffering from PD are unable to remain still without medication (Rusconi & Mitchener-Nissen, 2013). Additionally, the environment that an MRI presents is quite cramped, which might provoke feelings of claustrophobia in people. Even if the people could be forced inside or forced to lay still the resulting changes in neural activity might confound the results (Ganis et al., 2011; Rusconi & Mitchener-Nissen, 2013).

A basic fact of fMRI lie-detection is that the BOLD signal cannot directly show that a person is lying. Rather what needs to be done is to find the difference in activation that occurs during truthful responses and lies. Different activation has been found, especially in specific regions in the prefrontal cortex (Abe, 2011; Christ et al., 2009), and the idea is that one can correlate increased prefrontal activity to deception. By conducting experiments with low ecological, but high internal, validity the researchers tries to find the basis for deception, and what cognitive and neural regions are used during it. So far, no single region has been clearly correlated with deception, rather it has been found that most kinds of deception utilize regions related to working memory, inhibitory control, and cognition (Abe, 2011; Badre & Wagner, 2007; Christ et al., 2009; Tsuchida & Fellows, 2013). It should be noted that different types of deception will utilize different regions to a greater or lesser degree and that the activation pattern will also differ on an individual level. Thus, it is important to study deception on both a group and individual level (Rusconi & Mitchener-Nissen, 2013). Work is being done to determine the effects of emotions of deception and the effect of the salience of a stimulus on neural activation in relation to deception (Ito et al., 2011; Lee et al., 2010). This is very important as a common questioning technique is GKT, which essentially causes greater activation on more salient stimuli. This fact allows countermeasures but also requires the questions to be balanced so that a neutral stimulus does not cause the same activation as something related to a crime (which only the perpetrator would know) would. The balancing is important in both scientific as well as possible legal situations.

Non-neurological Limitations of Lie-detection

With the development of new technologies and techniques ethical, legal, and philosophical questions will often arise. Some are tied to the methodology being used, and can be handled in that arena. On the other hand, others require the involvement of philosophers and legal experts. The severity of these questions varies, but at the very least they need to be considered. Some issues are inherent to any lie-detection technology, but other are specific for fMRI lie-detection. For example, if fMRI is implemented in court, how should the evidence be handled? What laws protect the one being scanned and to what degree does he have the freedom to not be scanned? The issues are extensive and for reasons of space, only some will be touched on here.

Technology develops much quicker than laws can change and thus the introduction of new technologies can pose challenges to existing and new laws. Lie-detection is one of many technologies that challenges our laws and ethics, and Mitchener-Nissen (2013) describes eight categories that concern security technologies in general: “the causing of physical and mental harm, questions of legality, financial costs, liberties and human rights issues, broader public responses, issues of functionality, security and safety issues, and abuse/misuse issues” (Mitchener-Nissen, 2013, p. 2). The issues of harm as well as security and safety are not overly complicated in the case of fMRI lie-detection (though not negligible), and it might be too early to discuss how financial costs should be distributed. The functionality of fMRI lie-detection is a concern, but also something that is currently being studied and improved upon (as has been seen in this essay). Legality will need to be handled separately for each country, but some have pointed out that this is not be a matter for scientists to decide on (Schauer, 2010). Even though fMRI lie-detection has met much resistance in the scientific community (Rusconi & Mitchener-Nissen, 2013) the possible implementation of the technique in court will be handled by lawyers and legal experts (Schauer, 2010). There has been much discussion within the scientific community whether it is right to implement fMRI lie-detection or not, but Schauer's point is that scientists will only have an advisory role, the decision, by its nature, will be made by lawyers and legal experts.

Being such a new technique very few countries have produced any policies or laws concerning its possible implementation. A forerunner in this area is France, which has produced a legal article which states that the use of brain imaging techniques be limited to "medical purposes, scientific research, and in the context of judicial enquiries carried out by experts" (Rusconi & Mitchener-Nissen, 2013, p. 8). This was passed as a law in 2011 and bans imaging techniques from being used for commercial purposes. Some scientists have however criticized the law because "none of the neuroscientists consulted during the drafting process said that it [the use of imaging techniques in court] should be encouraged" (Oullier, 2012, p. 7). The most important part of the law is however that it expresses that informed consent must be obtained before any kind of imaging, and that this consent is revocable at any time (Rusconi & Mitchener-Nissen, 2013). Related to consent is the philosophical and legal issue of privacy. Though the proposal that fMRI scanning would be equal to mind reading might be exaggerated, the extent of the invasion of privacy still needs to be discussed. The clearest issue is that of the right to not self-incriminate. Unless fMRI scans would be deemed to be similar to physiological evidence such as DNA or blood analysis a person would be forced to incriminate oneself in a crime, against one's will. The right to not self-incriminate has even been argued, in the case of Funke v. France (1993), to be an implicit component of "one's *Right to a fair trial* under Article 6" (Rusconi & Mitchener-Nissen, 2013, p. 8) of the European Convention on Human Rights.

Discussion

Lie-detection using fMRI and the neuroscience of deception in general, have many interesting aspects to discuss. However, the most important aspect is clearly to discuss how the studies can have such differing designs and still often show similar activation patterns. Considering the difficulty of constructing ecologically valid experiments, to what degree are the results generalizable, and do they truly show the brain activity of deception?

In this essay the current state of the neuroscience of deception and many of the limitations as well as some of the strengths of fMRI lie-detection have been presented. The methodological issues that have been shown to exist are problematic and will often be hard to separate from other legal and philosophical issues.

A fundamental problem with a process such as deception is that it consists of many different cognitive processes and different types of deception, most likely, require a specific set. The most common regions found are VLPFC, DLPFC, APFC, and ACC (Abe, 2011; Christ et al., 2009), but are these regions vital to deception or to another process which may be close to deception but neither necessary nor dependent on it? According to Tsuchida and Fellows (2013) VLPFC, DLPFC, and ACC have all been proposed to function as general-purpose regions. On the other hand, specific functions have been connected to these regions as well. For example, activity in VLPFC seem to correlate with processes such as: task switching, reversal learning, response inhibition (Abe, 2011), semantic & episodic retrieval, post-retrieval selection, and rule-guided behavior (Badre & Wagner, 2007). These processes may be necessary for deception, but it might as well be so that the processes are entirely or partly independent. Semantic retrieval may be necessary for certain types of responses common during deception, but is semantic retrieval necessary for the deception to occur? To test this one would need to construct experiments focusing on specific types of deception, make sure that the situation is ecologically valid, similar to how that type of deception would look in real life, but also have enough control to make certain that it is deception being studied and not any other cognitive process, such as social interaction. All these parts are problematic.

How do we differ between one type of deception and another? There are studies using simple fabrications connected to working memory (Vartanian et al., 2013), deceptive behavior in gambling (Sip et al., 2010), simple fabrications in connection to memory and distorted memory (Abe et al., 2008; Lee et al., 2002), and many other. In the current literature deception is seen as multifaceted, and functions in conjunction with many other types of processes, but is vaguely defined. One of the most basic deceptions, a simple fabrication of fact, is defined by most as deception, but by its nature deception encompasses many types of behavior that need not be defined as deception by everybody.

How do we construct an ecologically valid experiment, especially while retaining sufficient internal validity to be able to be confident in our results? Disregarding different types of deceptive behavior, deception is problematic as it is often used during social interaction, is subjectively defined, and seen as morally bad (especially when others are deceptive; Christ et al., 2009). Social interaction can be achieved in an fMRI scanner, but the interaction will often be artificial, and face-to-face interaction, especially verbal communication, is close to impossible. Most agree that deception is something morally wrong (Christ et al., 2009), but not in all cases, nor do we feel that the same type of behavior in ourselves is as wrong as when others are deceptive. As deception is seen as morally wrong, however, it is difficult to construct an ethical experiment while retaining high ecological validity by including high risk, high gain, and social condemnation if the deceiver is found out. It will most likely be impossible to construct an experiment using real-life deception (Christ et al., 2009), but studies using more ecologically valid designs have found different patterns of activation (compared to less ecologically valid studies), specifically activity in amygdala (Abe et al, 2007; Baumgartner et al., 2009). The problem is to know if this activity is deception-related or related to the social interaction in the experiments. Baumgartner et al. (2009) proposes that it is due to being confronted, not deception.

Differing designs is one of the aspects that makes the seemingly consistent results found in studies on deception hard to accept at face value. This is supported by the fact that replications within the research area are very low (e.g. see Kozel et al., 2005; Kozel et al., 2009). Using different designs and participants Jiang et al. (2013) and Abe et al. (2009) found deception to be correlated with activity in DLPFC, but with different outcomes. In Jiang et al. lower activity in DLPFC was correlated with high ability to deceive (for people with ASPD), while Abe et al (2009) report that people with PD have a lower ability to deceive, while still correlating the activity in DLPFC with deception. Similarly, stimulation of DLPFC has been found to increase response time (Mameli et al, 2010) while others have not found this (Luber et al., 2009).

Complicating the case further, there is a wide variety in how scientists report the foci found. Some use anatomically defined regions while others use regions defined by function. Again some report Brodmann areas and some do not (Christ et al., 2009). Though this may be seen as a problem for research using brain scanning overall, the neuroscience of deception is affected by it as well.

The neuroscience of deception is limited by the physical limits of fMRI, such as not allowing certain people to be scanned. The research area can however use other types of methods, such as PET or tDCS, that have other strengths and limits, for example not necessarily requiring the participant to be still for a long time, allowing metal in or on the person, and pregnant women to be participants. If fMRI is to be used as a lie-detection technique, however, these limits are quite severe. As pointed out by Rusconi & Mitchener-Nissen (2013) we need to find the percentage of the population that cannot be scanned in an fMRI. Even if the percentage is found, and found to be acceptably low, the fact will remain that fMRI lie-detection will (at least in its current state) require the person being scanned to be compliant. While the use of fMRI lie-detection can take many forms, if it is to be used on suspects of a crime and not only in defense of the suspect the current technology is problematic. This is further complicated by that despite early assumptions, fMRI lie-detection has been found to be vulnerable to countermeasures (Ganis et al., 2011). Additionally, though similar activation was found in people with ASPD (Jiang et al., 2013), a diagnose in which deception is common, the activity was lower, indicating that the more used one is to lying, the harder it is to detect the deception through brain imaging.

Studying the neuroscience of deception has its hurdles, as has been shown in this essay, but to utilize fMRI for lie-detection raises several other issues of scientific, ethical, philosophical, and legal nature. Most scientists working in the area agree that fMRI lie-detection shows great promise, but at its' current state it is too unreliable and too little research has been done (Langleben & Moriarty, 2013; Rusconi & Mitchener-Nissen, 2013). Some argue for more caution (Abe, 2011), while others discuss the importance of an open dialogue, as the technology already exists and is being used (Garland & Glimcher, 2006; Langleben, 2008). As pointed out by Schauer (2010) whether and how fMRI lie-detection will be used in and around the court will ultimately be decided not by scientists, but by legal experts. Oullier (2012) illustrates this by saying that no scientist wanted France to encourage brain scans in court, yet it did. Even though the decision may be ultimately out of the scientists' hands, it makes informing both the legal experts and the public of the validity and reliability of fMRI lie-detection in its current state even more important. After all, the court is but one arena in which fMRI lie-detection can appear and unless a major breakthrough happens it seems that current fMRI lie-detection is not as good as it should be, if it is to be used. On the other hand, with a lie-detection rate of 47 % (Bond & DePaulo, 2006) neither are we. To entirely disregard fMRI lie-detection may be just as bad. Langleben and Moriarty (2013) hopes that with time and funding fMRI lie-detection will become as qualified as eyewitness testimony and states that forensic evidence that is currently common, such as fingerprints, also suffer greatly when it comes to validity. In an interesting article Faulkes (2011) argues that it would be ethically wrong to not implement fMRI lie-detection (when it has become better) as it could function as an argument against torture. In extreme situations torture is still being used, even though inhumane and morally condemnable. If fMRI lie-detection could replace it in the future it seems wrong to not support its development.

References

Abe, N. (2011). How the brain shapes deception: An integrated review of the literature. *The Neuroscientist*, *17*(5), 560–574. doi:10.1177/1073858410393359

Abe, N., Fujii, T., Hirayama, K., Takeda, A., Hosokai, Y., Ishioka, T., … Mori, E. (2009). Do parkinsonian patients have trouble telling lies? The neurobiological basis of deceptive behaviour. *Brain*, *132*(5), 1386–1395. doi:10.1093/brain/awp052

Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E., … Fujii, T. (2008). Neural Correlates of True Memory, False Memory, and Deception. *Cerebral Cortex*, *18*(12), 2811–2819. doi:10.1093/cercor/bhn037

Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience*, *19*(2), 287–295.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*(13), 2883–2901. doi:10.1016/j.neuropsychologia.2007.06.015

Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, *64*(5), 756–770. doi:10.1016/j.neuron.2009.11.017

Bennett, C. M., Baird, A. A., Miller, M. B., & Wolford, G. L. (2010). Neural correlates of interspecies perspective taking in the post-mortem atlantic salmon: An argument for proper multiple comparisons correction. *Journal of Serendipitous and Unexpected Results, 1*(1), 1-5.

Bhatt, S., Mbwana, J., Adeyemo, A., Sawyer, A., Hailu, A., & VanMeter, J. (2009). Lying about facial recognition: An fMRI study. *Brain and Cognition*, *69*(2), 382–390. doi:10.1016/j.bandc.2008.08.033

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214–234.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex*, *19*(7), 1557–1566. doi:10.1093/cercor/bhn189

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., … Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, *28*(3), 663–668. doi:10.1016/j.neuroimage.2005.08.009

Farah, M. J. (2014). Brain images, babies, and bathwater: Critiquing critiques of functional neuroimaging. *Hastings Center Report*, *44*(s2), S19–S30.

Faulkes, Z. (2011). Can brain imaging replace torture. *Global Virtue Ethics Review*, *6*(2), 55–78.

Ford, E. B. (2006). Lie detection: Historical, neuropsychiatric and legal dimensions. *International Journal of Law and Psychiatry*, *29*(3), 159–177. doi:10.1016/j.ijlp.2005.07.001

Fullam, R. S., McKie, S., & Dolan, M. C. (2009). Psychopathic traits and deception: Functional magnetic resonance imaging study. *The British Journal of Psychiatry*, *194*(3), 229–235. doi:10.1192/bjp.bp.108.053199

Funke v. France, 1 C.M.L.R 897 (ECHR 1993).

Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: Covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *NeuroImage*, *55*(1), 312–319. doi:10.1016/j.neuroimage.2010.11.025

Garland, B., & Glimcher, P. W. (2006). Cognitive neuroscience and the law. *Current Opinion in Neurobiology*, *16*(2), 130–134. doi:10.1016/j.conb.2006.03.011

Gracco, V. L., Tremblay, P., & Pike, B. (2005). Imaging speech production using fMRI. *NeuroImage*, *26*(1), 294–301. doi:10.1016/j.neuroimage.2005.01.033

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences*, *106*(30), 12506–12511.

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

Ito, A., Abe, N., Fujii, T., Ueno, A., Koseki, Y., Hashimoto, R., … Mori, E. (2011). The role of the dorsolateral prefrontal cortex in deception when remembering neutral and emotional events. *Neuroscience Research*, *69*(2), 121–128. doi:10.1016/j.neures.2010.11.001

Jiang, W., Liu, H., Liao, J., Ma, X., Rong, P., Tang, Y., & Wang, W. (2013). A functional MRI study of deception among offenders with antisocial personality disorders. *Neuroscience*, *244*, 90–98. doi:10.1016/j.neuroscience.2013.03.055

Kozel, F. A., Johnson, K. A., Grenesko, E. L., Laken, S. J., Kose, S., Lu, X., … George, M. S. (2009). Functional MRI detection of deception after committing a mock sabotage crime*. *Journal of Forensic Sciences*, *54*(1), 220–231. doi:10.1111/j.1556-4029.2008.00927.x

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, *58*(8), 605–613. doi:10.1016/j.biopsych.2005.07.040

Lagemann, T., Wolf, M., Ritter, D., Doucette, S., von Kummer, R., & Lewitzka, U. (2012). Cingulate cortex aplasia and callosal dysgenesia combined with schizencephaly in a patient with chronic lying. *General Hospital Psychiatry*, *34*(3), 320.e11–320.e13. doi:10.1016/j.genhosppsych.2011.10.002

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

Langleben, D. D., Loughead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, *26*(4), 262–272. doi:10.1002/hbm.20191

Langleben, D. D., & Moriarty, J. C. (2013). Using brain imaging for lie detection: Where science, law, and policy collide. *Psychology, Public Policy, and Law*, *19*(2), 222–234. doi:10.1037/a0028841

Lee, T. M., Liu, H.-L., Tan, L.-H., Chan, C. C. H., Mahankali, S., Feng, C.-M., … Gao, J.-H. (2002). Lie detection by functional magnetic resonance imaging. *Human Brain Mapping*, *15*(3), 157–164. doi:10.1002/hbm.10020

Lee, T. M., Lee, T. M., Raine, A., & Chan, C. C. (2010). Lying about the valence of affective pictures: an fMRI study. *PloS One*, *5*(8), e12291.

Luber, B., Fisher, C., Appelbaum, P. S., Ploesser, M., & Lisanby, S. H. (2009). Non-invasive brain stimulation in the detection of deception: Scientific challenges and ethical consequences. *Behavioral Sciences & the Law*, *27*(2), 191–208. doi:10.1002/bsl.860

Mameli, F., Mrakic-Sposta, S., Vergari, M., Fumagalli, M., Macis, M., Ferrucci, R., … Priori, A. (2010). Dorsolateral prefrontal cortex specifically processes general – but not personal – knowledge deception: Multiple brain networks for lying. *Behavioural Brain Research*, *211*(2), 164–168. doi:10.1016/j.bbr.2010.03.024

McCabe, D. P., Castel, A. D., & Rhodes, M. G. (2011). The Influence of fMRI Lie Detection Evidence on juror decision-making: fMRI evidence. *Behavioral Sciences & the Law*, *29*(4), 566–577. doi:10.1002/bsl.993

Mitchener-Nissen, T. (2013). Addressing social resistance in emerging security technologies. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00483

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

Rusconi, E., & Mitchener-Nissen, T. (2013). Prospects of functional magnetic resonance imaging as lie detector. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00594

Schauer, F. (2010). Neuroscience, lie-detection, and the law. *Trends in Cognitive Sciences*, *14*(3), 101–103. doi:10.1016/j.tics.2009.12.004

Sellal, F., Chevalier, Y., & Collard, M. (1993). “Pinocchio syndrome”: A peculiar form of reflex epilepsy? *Journal of Neurology, Neurosurgery, and Psychiatry, 56*, 936.

Sip, K. E., Lynge, M., Wallentin, M., McGregor, W. B., Frith, C. D., & Roepstorff, A. (2010). The production and detection of deception in an interactive game. *Neuropsychologia*, *48*(12), 3619–3626. doi:10.1016/j.neuropsychologia.2010.08.013

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, *12*(2), 48–53. doi:10.1016/j.tics.2007.11.008

Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Ying, Z., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport*, *12*(13), 2849–2853.

Spence, S. A., Kaylor-Hughes, C. J., Brook, M. L., Lankappa, S. T., & Wilkinson, I. D. (2008) ”Munchausen”s syndrome by proxy' or a ”miscarriage of justice”? An initial application of functional neuroimaging to the question of guilt versus innocence. *European Psychiatry, 23*(4), 309-314. doi: 10.1016/j.eurpsy.2007.09.001

Spence, S. A., Kaylor-Hughes, C. J., Farrow, T. F. D., & Wilkinson, I. D. (2008). Speaking of secrets and lies: The contribution of ventrolateral prefrontal cortex to vocal deception. *NeuroImage*, *40*(3), 1411–1418. doi:10.1016/j.neuroimage.2008.01.035

Thomas, B. C., Croft, K. E., & Tranel, D. (2011). Harming kin to save strangers: Further evidence for abnormally utilitarian moral judgments after ventromedial prefrontal damage. *Journal of Cognitive Neuroscience*, *23*(9), 2186–2196.

Tsuchida, A., & Fellows, L. K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex*, *49*(7), 1790–1800. doi:10.1016/j.cortex.2012.10.014

Vartanian, O., Kwantes, P. J., Mandel, D. R., Bouak, F., Nakashima, A., Smith, I., & Lam, Q. (2013). Right inferior frontal gyrus activation as a neural marker of successful lying. *Frontiers in Human Neuroscience*, *7*, 1-6. doi:10.3389/fnhum.2013.00616

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspectives on Psychological Science*, *4*(3), 274–290. doi:10.1111/j.1745-6924.2009.01125.x

Yang, Y., Raine, A., Narr, K. L., Lencz, T., LaCASSE, L., Colletti, P., & Toga, A. W. (2007). Localisation of increased prefrontal white matter in pathological liars. *The British Journal of Psychiatry*, *190*(2), 174–175. doi:10.1192/bjp.bp.106.025056