# Brain-based mind reading in forensic psychiatry: exploring possibilities and perils

Gerben Meynen[1,2,*]

1. Tilburg Law School, Tilburg University, Prof. Cobbenhagenlaan 221, 5037 DE Tilburg, The Netherlands
2. Faculty of Humanities, Department of Philosophy, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
*Corresponding author. E-mail: g.meynen@uvt.nl

## ABSTRACT

One of the areas in which brain-based mind reading (BMR) may be applied is forensic psychiatry. The purpose of this paper is to identify opportunities and challenges for forensic psychiatry regarding BMR. In order to do so, a conceptual framework for BMR will be introduced, which distinguishes between three basic types of BMR, based on how they relate to the subject's knowledge. In addition, three features of BMR techniques will be articulated: first, whether they require passive cooperation; second, whether they require active cooperation; and third, whether they require that the subject is awake. Each of the types of BMR entails specific chances and risks for forensic psychiatry, involving, for example, confidentiality in the doctor–patient relationship and the possibility of coercive use of BMR techniques. It will be concluded that apart from legal considerations, such as tests of admissibility of evidence, professional ethics is highly relevant.

KEYWORDS: brain imaging, forensic psychiatry, legal insanity, lie detection, mind reading, neurolaw

* Gerben Meynen, MD, is an endowed Professor of Forensic Psychiatry at Tilburg Law School, Tilburg University, and an endowed professor of Ethics and Psychiatry at the Department of Philosophy, Faculty of Humanities, VU University Amsterdam.

## 1. INTRODUCTION

Brain-based mind reading (BMR) could contribute to forensic psychiatric assessments of defendants, now or in the near future.[1] Yet, the use of mind-reading[2] techniques, broadly construed, has also been met with reserve, criticism, and skepticism. For instance, Pardo and Patterson (2013) argue that brain based lie-detection research may fail to detect what it aims to study: lies. They point to the difficulty of lying in a research setting in which lies are actually 'permitted' or even encouraged because they are required for performing the study. In such a context, they argue, it is impossible to *really* lie.[3] If research on brain-based lie detection would face such a fundamental problem, the application of such a technique in the court of law would lack a solid scientific basis. Apart from this research-related issue, ethical and legal qualms have been raised concerning the possible use of BMR against a person's or defendant's will.[4] Furthermore, there are technical concerns. For instance, even if a BMR technique would work in research settings, using 'willing' test subjects, actual defendants could take 'counter measures' to hinder or distort mind-reading procedures—which the technique itself may not be able to register—leading to false outcomes.[5]

Still, in principle, if the issues raised are properly dealt with, BMR techniques could be used to assess defendants, prisoners, and possibly also prospective jurors.[6] In this paper, I aim to explore major opportunities and challenges for future forensic psychiatric use of BMR techniques.[7] In order to structure the discussion, I will introduce and apply a tripartite conceptual framework of BMR techniques (see section 3). The framework will enable us to identify some BMR challenges particularly relevant for forensic psychiatry; these are mainly related to confidentiality and trust in the doctor–patient relationship and to coercive use of the techniques.[8] The purpose of the paper is, in other words, to bring together those elements of the literature on BMR that

---

1    Gerben Meynen, *Neurolaw: Recognizing Opportunities and Challenges for Psychiatry*, 41 J. Psychiatry & Neurosci. 3–5 (2016); Giuseppe Sartori, Silvia Pellegrini, & Andrea Mechelli, *Forensic Neurosciences: From Basic Research to Applications and Pitfalls*, 24 Curr. Opin. Neurol. 371–7 (2011); Michael S. Pardo & Dennis Patterson, Minds, Brains, and Law. The Conceptual Foundations of Law and Neuroscience (2013).

2    Some may prefer the term 'brain reading'. John-Dylan Haynes writes: 'Recently new brain imaging technology has been developed which one day might make it possible to read a person's thoughts directly from their brain activity with a high degree of accuracy. This novel approach in neuroscience is often referred to as 'brain reading' or, more technically, the "decoding of mental states"'. John-Dylan Haynes, Brain Reading, in I Know What you're Thinking: Brain Imaging and Mental Privacy 29–40 (Sarah Richmond, Geraint Rees, & Sarah Edwards eds., 2012).

3    Pardo & Patterson, *supra* note 1; Henry T. Greely, Mind Reading, Neuroscience, and the Law, in A Primer on Criminal Law and Neuroscience. A Contribution to the Law and Neuroscience Project, Supported by the MacArthur Foundation(Stephen J. Morse & Adina L. Roskies eds., 2013); Haynes, *supra* note 2. See also Gerben Meynen, Legal Insanity: Explorations in Psychiatry, Law, and Ethics, in International Library of Ethics, Law, and the New Medicine (David N. Weisstub & Dennis R. Cooley, eds., 2016).

4    Nita A. Farahany, *Incriminating Thoughts*, 64 Stan. L. Rev. 351 (2012); Pardo & Patterson, *supra* note 1.

5    Stephen J. Morse & Adina L. Roskies, eds., A Primer on Criminal Law and Neuroscience. A Contribution of the Law and Neuroscience Project, Supported by the MacArthur Foundation (2013).

6    Greely, *supra* note 3.

7    On the topic of BMR and forensic psychiatry, see also, eg, J. Arturo Silva, *Forensic Psychiatry, Neuroscience, and the Law*, 37 J. Am. Acad. Psychiatry L. 489–502 (2009); David Linden, Overcoming Self-Report: Possibilities and Limitations of Brain Imaging In psychiatry, in I Know What You're Thinking: Brain Imaging and Mental Privacy 123–35 (Sarah Richmond, Geraint Rees, & Sarah Edwards eds., 2012).

8    Even though, in my view, this framework helps to bring forward these challenges, other frameworks or perspectives could well result in the identification of similar challenges.

are specifically relevant to forensic psychiatry and to analyse how they are relevant, using a new conceptual framework for reading people's minds.

## 2. THE RELEVANCE OF BMR FOR FORENSIC PSYCHIATRY

Why would BMR techniques be of special interest to forensic psychiatry? There are two related reasons. First, more than any other medical discipline, psychiatry has to rely on the patient's *own account*.[9] Surely, if a surgeon wants to find out whether a patient's leg is broken, she is likely to ask questions to the patient about what happened, whether the patient is in pain, etc. Yet, at some point an X-ray, perhaps a CT-scan will be made and, in the end, the diagnosis will be based on the X-ray or CT-scan, not on the history taking. The surgical patient's subjective account is not irrelevant, it *guides* the investigation, and has implications for the surgeon's approach to the patient, but for the diagnosis, ultimately, it is the imaging technique that counts. Yet, in psychiatry, the situation is significantly different compared to other areas of medicine, as David Linden points out:

> Psychiatry is the oddball of medical disciplines because it has almost no tests that positively aid the diagnosis of a particular disease. Where diagnostic tests are applied, such as neuroimaging, these normally serve to exclude other, 'organic' causes for the reported symptoms and observed behavioural abnormalities. Although behavioural changes feature prominently in the diagnosis of some mental disorders, particularly those with childhood onset (autism, attention deficit/hyperactivity disorder, conduct disorder), most classic psychiatric diseases are largely diagnosed on the basis of patients' self-report.[10]

Why is psychiatry different from other medical disciplines in this respect? Psychiatric disorders are mental disorders, in which *subjective experiences* are the core of the disorder and the diagnostic process. For instance, to diagnose depression, the doctor needs to know something about her patient's mood and feelings, eg whether she can still enjoy life, feel happy from time to time. Such questions refer to DSM-5[11] *core* features of depression. How do psychiatrists—and other health professionals—acquire *information about* a person's subjective state? Information about subjective states has to come, in principle, from the subject himself. That is one of the special features of subjective states.

Suppose that the patient with the broken leg would believe that his leg was broken, but that he lied about it. Would it make a difference for the diagnosis? Probably not, the X-ray provides the relevant information. Now suppose that a psychiatrist has to make an assessment of a person who doesn't want to be diagnosed as having a depression, even though he is suffering from a depression, fulfilling all DSM-5 criteria for depressive disorder. This may cause real problems for the psychiatrist because to a considerable extent she has to rely on the patient's first-person account of his own subjective

---

9    For the matter discussed in this section, see also Meynen, *supra* note 3.
10   Linden, *supra* note 7, at 123. Note, however, that this situation is not entirely unique for psychiatry. Diagnosing headaches in neurology may also rely on the person's subjective account—instead of on brain imaging evidence.
11   American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* (5th ed., 2013).

states.[12] Clearly, there may be valuable additional sources of information, eg behavioral signs, medical files, and—as far as forensic psychiatry is concerned—police files, but this does not nullify the central role of history taking in psychiatry. Consequently, as far as psychiatrists have to rely on first-person accounts, the diagnostic process is vulnerable to lying (and even some behavioral symptoms could be faked). As Linden (2012, p. 123) puts it: 'This opens up the possibility that someone might "fake" a psychiatric diagnosis by pretending to suffer from symptoms that are not actually present'. In conclusion, the patient's—or defendant's—own account plays a central role in psychiatric evaluations, including diagnosis; to the extent that psychiatrists have to rely on a patient's—or defendant's—own account, BMR could be useful in forensic psychiatric assessments.

While the first reason for the potential relevance of BMR for forensic psychiatry concerns the *psychiatry* component, the second reason concerns the *forensic* element: the intersection with the law, in particular criminal law. Lie detection may be used in different contexts, but *the* context in which it is used is the context of law. This is not surprising, because in legal cases, in particular criminal cases, people may be more *tempted to lie and deceive* than in usual social and medical interactions. A defendant may not only be tempted to lie to the police, the prosecutor and the judge, but also to the psychiatrist who evaluates that defendant within the context of a criminal case.[13] So, within the legal context in which forensic psychiatry operates, the reliance on subjective accounts—always a bit of a challenge for psychiatric diagnosis and evaluation—may become a more serious problem.

In sum, because of the central role of a person's subjective account in psychiatric assessments and because of the specific legal context (with an increased risk of lying and deception), BMR techniques could, in principle, be useful in forensic psychiatry. In this paper, I will further explore such a possible use. In order to discuss BMR and its relevance for forensic psychiatry, I will first introduce a tripartite conceptual framework of BMR. The distinction between the three types of BMR in this framework is made based on the way in which they relate to the knowledge of the person who is examined. Next, I will consider some of the specific features of each of the three types and discuss their relevance for psychiatry.

## 3. TYPES OF MIND-READING: A THEORETICAL FRAMEWORK

In order to explore the relevance and dangers of BMR in forensic psychiatry, it may be helpful to categorize the different ways in which a person's mind can be read.

A wide range of techniques can be categorized as mind-reading techniques. In a chapter on mind reading, the law, and neuroscience, Greely discusses detecting lies, detection of memory or recognition, detecting pain, detecting bias, and detecting consciousness.[14] The neuroscientific techniques involved may also differ considerably: the

---

[12]  Linden (2012, p. 123) writes: 'Thus, a person can receive a diagnosis of a serious psychiatric disease purely based on reported symptoms'.

[13]  Maartje Katzenbauer & Gerben Meynen, Insanity: Neurolaw and Forensic Psychiatry, in THE INSANITY DEFENSE. MULTIDISCIPLINARY VIEWS ON ITS HISTORY, TRENDS, AND CONTROVERSIES (Mark D. White ed. 2017).

[14]  Greely, *supra* note 3.

implicit association test (IAT)[15] is very different from P300 measurements[16], which are very different from functional magnetic resonance imaging (fMRI).[17] Clearly, BMR techniques can be categorized in various ways, for example, by the neuroscientific techniques used (eg X-rays or electrophysiology) or brain processes (eg electrical activity of the neurons or blood flow) involved. Notably, in the *future*, additional methods and techniques are very likely to be developed. In fact, at present almost no technique appears to be ready for use in forensic psychiatric evaluations. Therefore, the topic of BMR basically derives its relevance and urgency from *anticipated* developments in the (near) future. And at present we do not really know the exact nature of the techniques that will eventually be ready for forensic psychiatric use. Therefore, it may be wise not to categorize BMR techniques based on (current) technical or neuroscientific features, but rather on distinctive *conceptual features*, which are likely to remain relevant for new technologies and neuroscientific approaches. I will follow this theoretical approach and thus distinguish between three types of BMR.

The tripartite framework that I introduce is meant to be simple, and broad and flexible enough to encompass very different mind-reading techniques. It should, in addition, enable us to explore some relevant issues for forensic psychiatric use of BMR techniques. The distinctive feature concerns the *subject's knowledge*. I will briefly discuss the three types of mind reading that can be theoretically distinguished based on how they relate to a subject's knowledge. Next, I will consider three additional issues that are relevant to the applicability of BMR techniques in each of the three categories.

The first type of mind reading makes use of an *external information source* to which the person's response is related. The subject knows something about the external source of information, and what he knows about this, is detected (eg showing images and recording a P300 signal, see section 3 for explanation). The second type detects phenomena a person knows, but there is no external information source: aspects of a person's mind are directly 'read', not the responses to an external stimulus (eg real-time recording of a subject's thoughts[18]). The third type detects information that *need not to be known* to the subject himself (eg IAT: people perform a task and based on their (unconscious) biases they may be slower or faster).[19] I will elaborate on these three types below.

---

[15] The IAT procedure consists of a series of tasks, eg, categorizing certain words that may be related with gender (if one targets gender-related attitudes), designed to 'measure implicit attitudes by measuring their underlying automatic evaluation. The IAT is therefore similar in intent to cognitive priming procedures for measuring automatic affect or attitude'. Anthony G. Greenwald, Debbie E. McGhee, & Jordan L. Schwartz, *Measuring Individual Differences in Implicit Cognition: The Implicit Association Test*, 74 J. Pers. Soc. Psychol. 1464–80 (1998).

[16] A P300 signal is an event-related potential component that is often considered to be associated with the *salience* of a stimulus for the subject. In certain experimental designs, it is, more specifically, considered an indication of recognition. See Section 3 for more on this technique.

[17] Techniques used for brain imaging include computed tomography (CT), magnetic resonance imaging (MRI), voxel-based morphometry, electroencephalography, magnetoencephalography, positron emission tomography, single photon emission CT, BOLD-fMRI, and near infrared spectroscopy. Sarah Richmond, Introduction, in I Know What you're Thinking: Brain Imaging and Mental privacy 1–10(Sarah Richmond, Geraint Rees, & Sarah Edwards eds., 2012).

[18] On real-time fMRI and the possibility of reading brain states, see, eg R. Christopher deCharms, *Applications of Real-Time fMRI*, 9 Nat. Rev. Neurosci. 720–9 (2008).

[19] This may be considered unconscious 'information'—for a critical discussion of the use of the term information in such a context, see Pardo & Patterson, *supra* note 1.

As is clear, *knowledge* is a central concept in the theoretical framework just described. There are, meanwhile, as Steup writes, 'various kinds of knowledge: knowing how to do something (for example, how to ride a bicycle), knowing someone in person, and knowing a place or a city'.[20] Regarding propositions, knowledge is often defined as justified true belief.[21] Yet, within the context of the present paper, such philosophical definitions are likely to be too strict and narrow. Therefore, in this paper, knowledge will be used more loosely, in the sense of 'being aware of something'. So, in this sense, a person may 'know' that he likes ice cream: he is aware of it—whether or not it is a *justified* belief is less important for our purposes. I will look in some more detail at the three types of BMR.

*Type I* refers to BMR techniques that relate a certain brain reaction to externally accessible information. The subject knows something *about* something that is externally presented. It is the brain response *in combination with the externally presented information* that is assessed. I give two examples of type I BMR:

1. *Lie detection*: part of the information is derived from outside sources: written or oral statements the person makes. The subject's mind is read merely to determine the truthfulness of the subject *as related to* the statement that is made. The technique may be directly brain based, or rely on assessing other parameters, outside of the skull, such as skin conductance.[22] Lie detection can only take place if the defendant makes relevant *statements* (that can be either truthful or deceitful). Note that, if (the subject believes that) the technique works well, the subject *knows* what the examiner knows about him or her: whether or not the subject has lied about a particular statement.[23]

2. *P300*: in a standard case, the subject looks at various images, and—at least in the criminal context—some of these pictures are related to a particular crime. For instance, the car used for escape, bags that have been used to put stolen money in. Ideally, if the subject sees random cars and bags, no P300 signal will be generated and recorded. However, when a picture of a meaningful, salient object is shown, the subject's brain may respond generating a P300 signal.[24] Note that such a response is very fast, and that it is not willfully generated. If a suspect responds time after time with a P300 signal to the pictures showing objects related to the crime, while not responding to other pictures unrelated to the crime, it becomes increasingly likely that he or she knows more about the

---

[20] Matthias Steup, *Entry Epistemology*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY. http://plato.stanford.edu/entries/epistemology/#JTB (accessed Feb 28, 2017).

[21] Matthias Steup, *Entry Epistemology*, Entry Epistemology, STANFORD ENCYCLOPEDIA OF PHILOSOPHY. http://plato.stanford.edu/entries/epistemology/#JTB (accessed Feb 28, 2017).

[22] On basic research concerning brain-based lie-detection using fMRI, see, for example, Daniel D. Langleben et al., '*Telling Truth from Lie in Individual Subjects With Fast Event-Related fMRI*, 26 HUM. BRAIN MAPP. 262–72 (2005). In 2005, they reported that lies were discriminated from truthful responses with an accuracy of 78%. See also Feroze B. Mohamed et al., *Brain Mapping of Deception and Truth Telling About an Ecologically Valid Situation: Functional Mr Imaging and Polygraph Investigation–Initial Experience*, 238 RADIOLOGY 679–88 (2006).

[23] I will only be considering lie detection, not possible use of some 'truth serum'.

[24] J. Peter Rosenfeld, P300 in Detecting Concealed Information, in xvii MEMORY DETECTION : THEORY AND APPLICATION OF THE CONCEALED INFORMATION TEST 319 (Bruno Verschuere, Gershon Ben-Shakhar, & Ewout Meijer eds., 2011).

crime.[25] Just like in the lie-detection case, the person knows what the examiner knows (assuming that the technique works fine).

Both examples of type I BMR we just discussed are 'one trick ponies': lie detection can only detect deceit, while P300 can only detect salience/recognition. In both techniques, it is the response *in combination with* the external stimulus (statement or images) that provides information about relevant aspects of a subject's state of mind (assuming the technique works well). Without the statements or pictures, we know nothing about the subjects mind. In fact, this is a *simple* form of mind reading, basically detecting a specific mental attitude or response regarding a specific phenomenon. Whatever else may be going on in a person's mind remains undetected. This implies that the examiner has to know what he or she is looking for, eg the truthfulness of a particular statement or the salience of a particular object (or series of objects).

*Type II* BMR does not involve an external information source. It detects the contents of the mind (or brain) as such. A requirement, however, is that the person knows—or is aware of—the information that is detected in the brain. All kinds of mental phenomena may be detected—'knowledge that', 'knowledge how', values, preferences, emotions, etc.—but the person has to be aware of these preferences, and has to know about them (which is in contrast to type III BMR). Just like in type I BMR, if the technique works well, the subject (defendant) therefore knows what the examiner knows about him or her. A further distinction between two subtypes of type II BMR can be made.

*Type IIa.* Subjective experiences of which the person is *aware at that moment*. So, the examiner actually knows what the subject is thinking about at that moment: such as: 'I like my ice cream' (when eating an ice cream), 'This is a beautiful painting' (when looking at Rembrandt's 'Night watch'), or whatever appears to be the proper reflection of such *real-time* conscious mental states. The work of Marcel Just and colleagues is relevant in this respect: students majoring in physics or engineering were asked to think of physics concepts such as gravity, electric current, and energy, while their brain activity was measured by fMRI. It turned out that 'individual concepts were identifiable from their fMRI signatures with a mean rank accuracy of .75'.[26] Earlier, Just et al. already succeeded in identifying emotions based on fMRI scanning.[27] Type IIa requires that the subject actually thinks about the issue. So, this is the 'I know what you are thinking' type of BMR, which may be normally associated with 'mind reading'.[28] In my view, some may consider this 'true' mind reading: a live recording of what is going on in a

---

[25] P300 could also be used for other interesting purposes: see on a P300 guided (virtual) wheelchair: Ulrich Hoffmann et al., *An Efficient P300-Based Brain-Computer Interface for Disabled Subjects*, 167 J. NEUROSCI. METHODS 115–25 (2008).

[26] Robert A. Mason & Marcel A. Just, *Neural Representations of Physics Concepts*, 27 PSYCHOL. SCI. 904–13 (2016).

[27] Karim S. Kassam et al., *Identifying Emotions on the Basis of Neural Activation*, 8 PLoS ONE e66032 (2013).

[28] For instance, a book on neuroscientific mind-reading is entitled 'I Know What You're Thinking: Brain imaging and mental privacy'. SARAH RICHMOND, GERAINT REES & SARAH J. L. EDWARDS, I KNOW WHAT YOU'RE THINKING BRAIN IMAGING AND MENTAL PRIVACY (2012), http://dx.doi.org/10.1093/acprof:oso/9780199596492.001.0001

person's mind. This type of BMR may also be used for consciousness detection and pain detection—which may be really valuable in courtroom cases.[29]

*Type IIb* is identical to type IIa; however, now the subject need not think about the issue at this particular moment, it is no real-time detection. The brain is approached as some 'information storage organ' and the information is just being retrieved from it. Humans have a lot of *knowledge* about numerous issues, which we just do not think about right now. Still, I know it, and I know that I know it. For instance, I know my mother's neighbor's name, even though I almost never think about it. This type of knowledge can be detected using type IIb BMR. Note that the scope of detecting knowledge in type IIb is much wider than in type IIa: only what the person is thinking about at that very moment can be detected in IIa.

In sum, both in type IIa and IIb BMR, the contents of what is read is 'epistemically accessible' to the subject. Another feature of type II BMR is that you do not have to determine a focus of the mind reading, as is required in type I mind reading. You just start the recording, and, at least in principle, you can just wait and see what will be detected. Notably, type II mind reading may also be highly valuable for people who are paralysed, and who have difficulty speaking, eg after a stroke.

*Type III* concerns 'non-knowledge' mind reading. This is similar to type II mind reading in the sense that no external source of information is required: it merely concerns information derived from the brain itself. Meanwhile, it is different from type II in that what is being read is not known to the subject, it is not something that is 'epistemically accessible' to the person. If the technique works well, the examiner has knowledge (derived from the subjects brain), the subject himself does not have. Therefore, the subject does not know what the examiner knows after the test. In a way, the examiner may know the subject better than the subject knows himself. One example of type III BMR would be detecting biases in prospective jurors (see Greely 2013 on detecting such biases, eg using the IAT). Note that people may not be aware of some of their biases. As Croskerry writes: 'First, many decision makers are unaware of their biases, in part because our psychological defense mechanisms prevent us from examining our thinking, motivation, and desires too closely'.[30] Consequently, if asked whether they are biased toward a certain group of people using lie detection (type I BMR) these people will pass with flying colors. When type II BMR is used, they may still find that the person actually thinks or believes: 'I am not biased'. Still, type III BMR may show unconscious brain processes revealing significant biases—which is important information about that person's mind.

In fact, the distinctive—and valuable—feature of type III BMR is that the measurement does *not require* the person to know, even though he may know. For example, for the IAT, it is not required that the person knows about his own biases, even though he may know them.[31] Another example of using type III BMR would be recovering a

---

[29] Greely, *supra* note 3; Adrian M. Owen, When Thoughts Become Actions: Neuroimaging in Non-Responsive Patients, in I KNOW WHAT YOU'RE THINKING: BRAIN IMAGING AND MENTAL PRIVACY 73–87 (Sarah Richmond, Geraint Rees, & Sarah Edwards eds., 2012).

[30] Pat Croskerry, *From Mindless to Mindful Practice—Cognitive Bias and Clinical Decision Making,* 368 NEW ENG. J. MED. 2445–8 (2013).

[31] See Greenwald et al. on IAT: 'Implicit attitudes are manifest as actions or judgments that are under the control of automatically activated evaluation, without the performer's awareness of that causation (…) A possible property of the IAT—and one that is similar to a major virtue of cognitive priming methods—is that it may

person's forgotten memories (which is not technically possible at present).[32] This concerns knowledge that is *no longer available*. For instance, the defendant once knew the name of the person who sold him a car, but now he has forgotten. Using type III BMR, the forgotten memory may somehow be 'retrieved' or 'reconstructed'. Another—and important—type III technique is neuroprediction.[33] The person himself may genuinely believe that he no longer poses a threat to society, and that he will never again commit a crime, but BMR may reveal a high risk of recidivism. Since in type III BMR, the subject may or may not be aware of the information that is yielded using the technique, he may or may not know what the evaluator knows after using it. In fact, the subject could have the false impression that he knows what the evaluator knows—because the evaluator may know the subject 'better than he knows himself'.

Type III BMR techniques may not even require that the person is awake. Suppose a victim of a violent crime is comatose (as a result of the crime). Using such a type III BMR technique, information about a perpetrator may still be obtained.[34] So, the comatose victim could nevertheless provide evidence as a 'witness' as a result of such BMR. While type II BMR may be valuable in people with difficulty speaking and writing (eg after a cerebrovascular accident or people with a locked in syndrome) but who are still awake, type III BMR may be used in people who are comatose (assuming that there is no awareness during the coma).

Note that, in principle, brain-based *diagnostic* procedures for mental abnormalities count as type III BMR—assuming that mental illness refers to mind states. Suppose that a person's character has changed over the past weeks or months. A brain scan may reveal a tumor; such a finding provides information about his mind—in a broad sense. For instance, based on the scan, we may be able to predict deterioration of the mental abnormalities if no surgical intervention takes place. So, in this sense, we know something about the subject's mind, the person need not know himself. We should also realize that, currently, in psychiatry such diagnostic techniques hardly exist: schizophrenia, depression, bipolar disorder, personality disorder, etc, are not diagnosed using brain imaging techniques. Nevertheless, brain scans are used to detect brain trauma, brain tumors, and neurodegenerative diseases, each of which may lead to mental disorders or disturbances. In some legal cases, brain scans of tumors, traumas, and neurodegenerative diseases have already

---

resist masking by self-presentation strategies. That is, the implicit association method may reveal attitudes and other automatic associations even for subjects who prefer not to express those attitudes'. Greenwald et al., *supra* note 15.

[32] Notably, ample research on memory suggests that our memories are reconstructed; Aya Ben-Yakov, Yadin Dudai, & Mark R. Mayford, *Memory Retrieval in Mice and Men*, 7 COLD SPRING HARB. PERSPECT. BIOL. (2015). Forgotten memories appear not to 'lie somewhere in the brain' as if the brain records information like a videotape. Yet, much is still unknown about memory and memory retrieval. Interestingly, Farah and Wolpe consider the possibility of distinguishing between true and false memories; Martha J. Farah and Paul Root Wolpe, *Monitoring and Manipulating Brain Function: New Neuroscience Technologies and Their Ethical Implications*, 34 HASTINGS CENT. REP. 35–45 (2004).

[33] Thomas Nadelhoffer et al., *Neuroprediction, Violence, and the Law: Setting the Stage*, 5 NEUROETHICS 67–99 (2012). See for prediction of behavior also Geraint Rees & Ryota Kanai, Predicting Human Behaviour from Brain Structure, in I KNOW WHAT YOU'RE THINKING: BRAIN IMAGING AND MENTAL PRIVACY 59–69 (Sarah Richmond, Geraint Rees, & Sarah Edwards eds., 2012). See for prediction of mental disorder, such as schizophrenia: Linden, *supra* note 7.

[34] On mind reading (awareness detection) in non-responsive patients, see Owen, *supra* note 29.

played a role, even though their interpretation may still be a puzzle, especially as far as the legal consequences are concerned.[35]

In sum, type I and II BMR both detect knowledge on the part of the subject. They require the subject to know something in order for the technique to be useful. Type III does not rely on or require knowledge on the part of the subject.

More can be said about the mind-reading techniques. Three additional issues are relevant (though not unrelated to the distinctions made above between the three types of BMR), also for medical use of these techniques.

First, usually, the *cooperation* will be required in the sense that the person has to be willing, eg to enter a scanner and to lie still. I will consider this *passive* cooperation: the main issue is that the person should *not resist* the measurements being performed.[36] Some (future) BMR techniques, meanwhile, may not depend on a person's passive cooperation; they could also take place when people physically resist the measurements. Of note, 'cooperation' in this sense does not necessarily mean that the subject does not try to manipulate or distort the test results.

Second, it may be required that the person performs a *certain (mental) task* in order to get the right recordings or measurements. For instance, P300 (type I BMR) evaluations require subjects to look at certain pictures; IAT (type III BMR) requires performing a task. I will consider this *active* cooperation. While the procedure takes place, the person has to *do* something in order to make the measurement work. Even a resting state measurement requires that the person actively 'rests'. Also regarding this second issue, 'cooperation' does not necessarily mean that the subject does not try to manipulate or distort the test results.

Third, in the cases in which no cooperation whatsoever is required, it might even be possible that the person is asleep, depending on the technique. Even though the 'mind' is read, at this very moment there may be no mental activity (unless the subject is dreaming). So, it is not required that the person is *awake*. This is relevant for subjects—for example, victims—who are comatose. In fact, this may also be very relevant for coercive use of BMR: the subject may be sedated before using a BMR technique. This is more likely to be possible in type III BMR, in which the procedure does not rely on what the person knows (eg a diagnostic procedure).

---

35  Gerben Meynen, Neuroethics of Criminal Responsibility. Mental Disorders Influencing Behavior, in THE ROUTLEDGE INTERNATIONAL HANDBOOK OF BIOSOCIAL CRIMINOLOGY 544–57 (Matt DeLisi & Michael G. Vaughn eds., 2015).

36  Yet, in principle, a person could also be scanned against his or her will, just like in certain cases fingerprints, blood, and DNA can be taken against a person's will. In addition, all subtypes in which active cooperation is required (so, in which a task has to be performed) require cooperation. Campbell and Eastman (2012, p. 150) write: 'The neuroscientific investigation of a defendant, or patient, is different from other types of investigation in that it is potentially achievable without co-operation from the subject (assuming sedation could be used, and where the particular test does not require cognitive co-operation). Hence, it may often be possible to pursue investigation in the absence of consent. The subject, therefore, can potentially lose control of information about their own neurobiology.' C. Campbell & N. Eastman, The Neurobiology of Violence: Science and Law, in I KNOW WHAT YOU'RE THINKING: BRAIN IMAGING AND MENTAL PRIVACY 139–53 (S. Richmond, G. Rees, & S. Edwards eds., 2012).

## 4. RELEVANCE OF THE THREE TYPES OF BMR FOR FORENSIC PSYCHIATRIC ASSESSMENTS

Already in 2008, Langleben and Dattilio wrote: 'Mind-reading with fMRI is scientifically more challenging and even less developed than lie detection, yet it is of great interest and relevance to *forensic psychiatry*. Despite the multitude of technical challenges, we believe that fMRI is a powerful technology that may, together with the more established structural MRI, form a discipline of forensic MRI in the foreseeable future.'[37] In this section, I explore the relevance of each of the three BMR types for forensic psychiatry, focusing on the assessment of defendants.

*Type I.* As discussed, the reliability of forensic psychiatric assessments of defendants is an issue.[38] The possibility of faking and malingering is a central topic regarding these assessments.[39] Suppose a defendant claims: 'Just before I attacked my neighbor, I heard a voice commanding me to stab him. I could not but obey this voice'. If the psychiatrist believes this statement, this may have major impact on his conclusion regarding the defendant's insanity (especially if a control prong is part of the legal test for insanity in that jurisdiction).[40] But it is known that 'command auditory hallucinations are easy to fabricate'.[41] Therefore, it may be helpful to test whether or not the defendant is lying when he claims that he heard such a commanding voice he could not but disobey. To be clear, even a defendant who does not lie, may somehow be mistaken. Still it could be very valuable to know whether the defendant is truthful or not—which could, in principle, be tested using type I BMR.

This is an example showing that type I mind reading may be useful in psychiatric evaluations of defendants. In fact, Grubin has pled for using lie-detection in forensic psychiatry:

> The use of polygraphy as a gimmick on daytime television and in a range of populist applications such as testing the fidelity of spouses should not obscure its potential benefits in forensic settings. It is employed widely and with success in the United States and in a large number of other countries in criminal justice and national security settings. So long as it is not seen as the equivalent of Wonder Woman's magic lasso and is always included as part of a larger package of information, its application to forensic psychiatry should be welcomed. Indeed, given what we know about the efficacy of polygraph testing with sex offenders, one might argue that it is no longer a question of why we should use it in forensic psychiatry, but why we don't.[42]

Nevertheless, current lie detection is controversial, as Simpson (2008) writes: 'The reliability and validity of the polygraph are controversial'. BMR lie detection is bound

---

[37] Daniel D. Langleben & Frank M. Dattilio, *Commentary: The Future of Forensic Functional Brain Imaging*, 36 J. AM. ACAD. PSYCHIATRY L. 502–4 (2008). Emphasis added.

[38] Grubin, *infra* note 42.

[39] Sartori et al., *supra* note 1.

[40] On this topic, see also Gerben Meynen, *A Neurolaw Perspective on Psychiatric Assessments of Criminal Responsibility: Decision-Making, Mental Disorder, and the Brain*, 36 INT. J. L. PSYCHIATRY 93–9 (2013); Meynen, *supra* note 3.

[41] Phillip J. Resnick & James Knoll, *Faking It: How to Detect Malingered Psychosis*, 4 CURR. PSYCHIATRY 13–25 (2005).

[42] Don Grubin, *The Polygraph and Forensic Psychiatry*, 38 J. AM. ACAD. PSYCHIATRY L. 446–51, 450 (2010).

to be controversial as well, at least during the first stages of development and forensic psychiatric use.

Simpson also points to the fact that it 'is important to bear in mind that, like the polygraph, fMRI lie detection requires a willing subject'.[43] So, the topic of cooperation—passive as well as active—discussed above, is highly relevant here. This may be especially true in jurisdictions where insanity is not a defense raised by the defendant, but instead court ordered (for example, in the Dutch legal system, forensic psychiatric evaluations can be court ordered).[44] If the defendant raises the defense himself, it appears likely that he will cooperate with the methods used in a psychiatric evaluation. However, when the evaluation is court ordered, the defendant may not at all agree with the evaluation, and in such cases of disagreement, the defendant's cooperation with the lie detection is less likely (as already mentioned, 'cooperation' in this sense does not necessarily mean that the subject does not try to manipulate the test results). Finally, we should recognize that a truthful statement need not be true. People may be mistaken, in particular in cases of dementia or other neurocognitive disorders. This means that psychiatrists will have to try to verify relevant statements. Inconsistencies should be further examined.[45]

*Type II* BMR could be helpful as well in forensic psychiatry. Consider, again, the commanding voice. A defendant may claim that he almost constantly hears commanding voices—that made him commit a terrible crime—and that he is actually hearing them right now. BMR of the type IIa could inform a psychiatrist about the actual, real-time experiences of that person. If the person hears such a voice at this moment, this phenomenon will be detected (if the technique works properly). First, this finding is helpful in diagnosing the defendant's condition; second, it provides information about the actual psychopathological phenomena the patient is experiencing. Note that if a BMR type II technique shows that a defendant is currently hearing voices, this does not prove that he heard them at the moment of the crime. Still, in a way, it appears to corroborate his statements.

*Type IIb* is clearly also relevant in forensic psychiatry. The defendant may not be willing to talk to the psychiatrist at all (refuse cooperation during a court-ordered psychiatric evaluation). In such a case, it could be helpful to use type IIb BMR to find 'information in the brain' regarding delusions, sexual interests, or aggressive impulses. The person may not be experiencing them right now (so type IIa BMR would not detect these phenomena), but in certain situations, the delusions, sexual interests, and aggressive impulses may be very relevant: in such situations, they may lead to a crime.

*Type III.* Forensic psychiatrists may well be interested in information 'stored in the brain'[46]—irrespective of whether or not the defendant knows about it. For instance,

---

43  J. R. Simpson, *infra* note 48, at 493 (2008).

44  On psychiatric evaluations of defendants in the Dutch legal system, see, eg, Gerben Meynen, Legal Insanity and Neurolaw in the Netherlands: Developments and Debates, in Legal Insanity and the Brain: Science, Law and European Courts (Sofia Moratti & Dennis Patterson eds., 2016).

45  In principle, although I do not see an immediate form of use in psychiatric evaluations of defendants, P300 measurements could also be helpful. Showing pictures of possibly familiar objects to people does not seem to serve a direct psychiatric purpose. Still, this may be found or developed, probably with some modifications of the technique.

46  Pardo & Patterson (2013) would not favor this way of putting it, but it is hard to find better phrasings.

certain memories that are forgotten or otherwise inaccessible could be relevant to retrieve. People suffering from memory deficits in Alzheimer's disease may not be able to recall certain experiences, but it may be valuable to reconstruct what they actually experienced at the time of the crime. Perhaps, from a Freudian perspective, one might argue that type III BMR could reveal a person's unconscious conflicts as well. Psychiatrists will also be interested in certain brain features that increase the chance of extreme anger that may lead to violence. The subject himself may not at all know about such things, but that does not diminish their forensic psychiatric significance. Psychiatrists could even examine a defendant who is in a comatose state after a crime. In general, psychiatrists may be very interested in aspects of a person's mind that are not known to the person himself. This is particularly true for predicting future dangerousness: what will happen in the future may not be clear to the subject, but still it is highly relevant from a forensic psychiatric point of view. Therefore, type III BMR could be very helpful: it may answer certain psychiatrically and legally relevant questions more directly than type II BMR.

Meanwhile, using type I and II BMR in forensic psychiatry could well harm the doctor–patient relationship. The reason is that using these techniques may be considered a strong sign of distrust regarding the subject. As made clear, in type I and II BMR the defendant who is evaluated already *knows* what the examiner wants to know. Yet, the reason to use type I and II techniques will often be that the examiner does not *believe or trust* the subject (or the person *does not want* to share the information). This is especially true in lie detection, but it may also be relevant for real-time BMR (type II). Creating a situation of distrust should in principle be avoided in forensic psychiatry, as a subspecialty in medicine, in which a relationship of trust is usually considered crucial. This is one of the reasons why, in my view, especially concerning *forensic psychiatry*, we should distinguish between type I and II BMR techniques on the one hand, and type III on the other, since in the latter case the patient himself need not have epistemic access to the issues of interest. In principle, using a type III BMR technique need not in itself express a similar level of distrust as lie detection, and therefore it does not have to undermine the doctor–patient/defendant relationship in the way a type I or II BMR techniques may. I emphasize this point within the context of forensic psychiatry, because a relationship of trust is usually much more relevant for a psychiatrist (who is a medical doctor and health care professional) than for, eg a police officer or a prosecutor or judge who might use a BMR technique.

Note that all three BMR types can help diagnose a mental disorder in a defendant in forensic psychiatry. Type I may be useful verifying the *truthfulness* of claims made by the defendant about psychopathological experiences, such as the claim that he hears voices talking to him. Type II may provide *direct* information about the defendant's *actual experiences*, such as the defendant experiencing voices talking to him at this very moment (which could, together with other information, lead to the diagnosis schizophrenia). Type III may be useful detecting *brain changes* that are known (not yet, but possibly in the future) to be *accompanied* by hearing voices.[47]

---

47 Some areas or circuits that may be involved have already been identified, Linden, *supra* note 7. Linden also uses the term 'neural marker' in this context.

## 5. CAVEATS AND PERILS

There is a general caveat that BMR techniques may not be as sound and adequate as we would wish, which has rightfully been articulated by various authors, for instance, from a philosophical perspective, by Tim Bayne:

> …identifying someone's mental states on the basis of information about their neural states is far from direct or unproblematic. 'Brainreading' is possible, but it is a risky business, for it requires a host of assumptions, many of which will be controversial.[48]

In fact, there are various general methodological issues concerning the use of neuroscience data in the courtroom, eg a lack of replication of neuroscientific findings, the small sample sizes used in the studies, as well as the intricacies of the interpretation of functional brain imaging.[49] In what follows, I will formulate some more specific qualms and caveats regarding each of the three types of BMR.

*Type I.* This type of BMR involves a real danger: that the simplicity of the BMR technique (at least compared to type II and III) leads to jumping to conclusions, because there is basically one stimulus and one response. So, if a certain response is observed after the stimulus, we may be seduced to immediately conclude that the response is a *response to* the stimulus. But it need not be. For instance, sexual arousal may arise from other sources than the specific stimulus presented. Or skin conductance may change due to other factors, not having to do with truthfulness of deceit.[50] So, even though the test may *suggest* simplicity, it may be highly complicated and—in fact—tricky.[51] Clearly, such a caveat also applies to forensic psychiatric use. In addition, there is the topic of 'distrust' related to these techniques, which could undermine a working alliance between the psychiatrist and the subject (see however below). Since active cooperation is required (at least in the type I techniques we discussed), real coercive use is not possible. Yet, offers may still be made that are hard to refuse for a subject (such offers may also be considered threats). This latter possibility may well entail both legal and ethical challenges for forensic psychiatrists.

*Type II.* If this type of mind reading would become possible, it is hard to fathom the consequences. It really brings up fundamental issues about a person's privacy, since

---

48  Tim Bayne, How to Read Minds, in I Know What You're Thinking: Brain Imaging and Mental Privacy 41–58 (Sarah Richmond, Geraint Rees, & Sarah Edwards eds, 2012). For some concerns from a forensic psychiatric perspective, see also Joseph R. Simpson, *Functional Mri Lie Detection: Too Good to Be True?*, 36 J. Am. Acad. Psychiatry L. 491–8 (2008); Sartori et al., *supra* note 1.; Silva, *supra* note 7. For general qualms regarding the use of neuroscience in the courtroom, see Owen D. Jones et al., *Neuroscientists in Court*, 14 Nat. Rev. Neurosci. 730–6 (2013).

49  Campbell & Eastman, *supra* note 36; Morse & Roskies, eds., *supra* note 5, on pitfalls in fMRI, see also Sven Haller & Andreas J. Bartsch, *Pitfalls in fMRI*, 19 Eur. Radiol. 2689–706 (2009).

50  As Wolpe et al. state, 'To create a test that truly measures verisimilitude or salience, the relation between the measured signal and the physiological chain of events coupling a behavior with the signal must be fully characterized.' Paul Root Wolpe, Kenneth R. Foster, & Daniel D. Langleben, *Emerging Neurotechnologies for Lie-Detection: Promises and Perils*, 5 Am. J. Bioethics 39–49 (2005).

51  Considering the issue of the 'illusory accuracy' of BMR in general, Farah and Wolpe state: 'Although brainwaves do not lie, neither do they tell the truth; they are simply measures of brain activity. Whether based on regional cerebral bloodflow or electrical activity, brain images must be interpreted like any other correlate of mental activity, behavioral or physiological. Brain images and waveforms give an impression of *concreteness and directness* compared to behavioral measures of psychological traits and states, and high-tech instrumentation lends an aura of accuracy and objectivity.' Farah and Wolpe, *supra* note 32(emphasis added).

now, the most private thing we 'have', is real-time accessible to others: our mind.[52]
This would be a new phenomenon for mankind. It will create some positive possibil-
ities, no doubt, but it will also confront us with fundamental ethical, legal, and societal
questions.[53] In forensic psychiatric practice, this type of mind reading may, in princi-
ple, be used *against a person's will*. The reason is that no active cooperation need be
required—at least as long as no task has to be performed the possibility of coercive use
has to be taken into account. So, more than type I, this type may raise the question for
forensic psychiatrists about coercive use.

Type IIb of mind reading is—in principle—able to detect *all* epistemically acces-
sible information (including attitudes, preferences, etc.), possibly in a very short time.
In terms of science fiction, a copy could be made of what a person knows—broadly
construed—possibly within seconds. Now, in the future, there may be purposes for
performing such a 'total mind scan', which is likely to require some 'big data' approach
in addition to a sophisticated neuroscientific technique. It should be emphasized that
there are not just legal issues relevant to the topic of privacy (such as the Fourth Amend-
ment in the USA), there is also the medical ethical issue of confidentiality. Even though
privacy and confidentiality[54] are generally considered relevant to BMR,[55] this topic has
specific relevance to forensic psychiatry. Confidentiality is a core characteristic of med-
ical practice; as a principle, medical professionals have a fundamental obligation to re-
spect the confidentiality of the medical—and therefore psychiatric—information they
obtain. The responsibilities for medical doctors in this respect are stricter than for many
other professionals in our society. Therefore, the question arises: to which extent should
forensic psychiatrists share information yielded by BMR type II with others in a crimi-
nal case? Type II BMR may reveal *a lot* of information.[56] Therefore, the psychiatrist will
have to make a decision about the relevance of certain information within the context of
his specific task. Confidentiality is a central medical ethics principle[57] (as well as health
law principle). It may be the case that type II BMR reveals information the psychiatrist

---

[52]  Daniel V. Meegan, *Neuroimaging Techniques for Memory Detection: Scientific, Ethical, and Legal Issues*, 8 Am. J.
    Bioethics 9–20 (2008); Richmond et al., *supra* note 28.

[53]  On such questions, see, eg, Wolpe et al., *supra* note 50, as well as comments to this paper in the *American Journal
    of Bioethics*. See also Daniel D. Langleben & Jane Campbell Moriarty, *Using Brain Imaging for Lie Detection:
    Where Science, Law and Research Policy Collide*, 19 Psychol. Pub. Pol.' L. 222–34 (2013). And Martha J. Farah
    et al., *Functional MRI-Based Lie Detection: Scientific and Societal Challenges*, 15 Nat. Rev. Neurosci. 1464–80
    (2014).

[54]  On the distinction between privacy and confidentiality, Fenton et al. write: 'Though issues of privacy and con-
    fidentiality often overlap, they can be usefully distinguished: privacy demands discretion in the collection of
    information about others; confidentiality demands respect for the wishes of persons about whom private in-
    formation has been collected as regards the management (eg storage, disclosure, and distribution) of that in-
    formation.' Andrew Fenton, Letitia Meynell, & Francoise Baylis, *Ethical Challenges and Interpretive Difficulties
    With Non-Clinical Applications of Pediatric fMRI*, 9 Am. J. Bioethics 3–13 (2009).

[55]  See Richmond et al., *supra* note 28 and Farah and Wolpe, *supra* note 32.

[56]  On the issue of confidentiality and real time (fMRI) mind reading, see deCharms, *supra* note 18: 'The foresee-
    able ability to read the state of a person's brain and thereby to know aspects of their private mental experience,
    as well as the potential to incorrectly interpret brain signals and draw spurious conclusions, raises important
    new questions for the nascent field of neuroethics. One can imagine the scenario of someone having their pri-
    vate thoughts read against their will, or having the contents of their mind used in ways of which they do not
    approve. Many of these concerns also apply to non-real-time applications'. In my view, it may even more rele-
    vant in type IIb BMR than in real-time (type IIa) BMR.

[57]  Laura Weiss Roberts, A Clinical Guide to Psychiatric Ethics (2016).

should not use in his report, as it is not relevant to the assignment, eg regarding certain aspects of sexual fantasies—thus treating them as 'confidential'. Yet, in other cases, such sexual fantasies are highly relevant to the assessment, and the psychiatrist will have to add them in his or her report. Type II and III BMR used by a forensic psychiatrist may even yield crucial information regarding *intent*, and issues related to the question *Who did it*? The reason that this may occur in type II and III BMR is that they may function as a 'trawl': they may pick up *all kinds of mental content indiscriminately*, whereas type I has exclusive focus because it relates to an outside source that limits the scope of the findings. How (not) to use the potential wealth of information yielded in this way will be a matter of serious legal and ethical considerations for forensic psychiatrists in the future. In other words, there is not just the legal issue of privacy involved in type II BMR, but, if used by a forensic psychiatrist, also medical or professional ethics—as well as health law. As Campbell and Eastman point out: 'Law represents a potential safeguard against abuse. However, the weaker the "probative" and "reliability" tests applied by law to expert neuroscientific evidence the greater the "room" for its misuse, including politically' (147–48). Forensic psychiatrists have an independent duty to consider their professional ethics obligations.

*Type III.* In type III BMR, one is not directly reading a person's mind in the sense of recording the things he is actually aware of or that are 'epistemically accessible' to him. In fact, as already made clear, some diagnostic procedures may count as type III mind reading, just as long as they relate to a (past, current, or future) state of mind. For instance, detecting a certain type of tumor in a particular area of the brain using MRI may enable a neurologist to say something about what a patient experiences now, or in the near future; he may be able to tell whether or not the person's visual experiences will be distorted. In such a case, 'standard' medical ethics applies. But in other cases, the issues discussed regarding type II may be relevant as well.

In general, *the* ethical caveat regarding all three types of BMR concerns the use of these techniques against a person's will, or within the context of an offer that a person cannot refuse.[58] Depending on the precise type of BMR, it may or may not be technically possible to perform recordings against a person's will. But in some cases BMR will be possible without the person having to cooperate, especially in type III BMR. In some type III BMR techniques, if the person would be made to sleep, using some anesthetic, the recordings could still take place. In cases where cooperation is required in the sense that a person has to perform a certain task, it may still be possible to force people to undergo a test[59] by making a severe threat.[60] It is a legal as well as ethical issue to what extent and in which circumstances such a use of BMR would be permissible. So, notably, the ethical and legal issues do not only arise in cases where no cooperation

---

[58] Compulsory use of BMR will, clearly, raise questions about the right against self-incrimination, the Fifth Amendment in the US context, Pardo & Patterson, *supra* note 1. At the same time, as Meegan (2008, p. 16) notes, 'if the examination is voluntary, the suspect should rightly be concerned about how the courts will perceive a refusal to be examined'.

[59] This possibility is also mentioned by Linden (2012, p. 134), for forensic settings.

[60] I have tacitly assumed that the techniques are non-invasive. Meanwhile, it could be that future BMR requires, eg, some contrast to be injected. This would open up another area of legal and medical ethical qualms.

is required, but also regarding situations in which people are threatened or in which 'offers are made that cannot be refused'.[61]

*The* general methodological caveat concerns the possibility of countermeasures and intentional distortion of the measurements.[62] As Haynes writes, regarding classical polygraph lie detection:

> Interestingly, polygraphy is quite reliable when applied to inexperienced subjects. The problem of classical polygraphy, however, is that arousal — which is used as a physiological marker of deception — can be affected by mental factors other than deception (such as general anxiety), and it can also be deliberately manipulated. For example, it has been repeatedly shown that subjects can deliberately and selectively control their level of arousal and thus distort polygraph tests… Instructions on how to do this are freely available on the internet. Therefore, where manipulation of polygraphy results by trained subjects cannot be excluded, the validity of the tests remains doubtful.[63]

The same may well be true for yet to be developed types of BMR. So, it may be the case that a certain BMR technique works fine, as long as the subject wants to make the best of it. But that the very same technique may lead to false results in case the subject uses countermeasures to distort the recordings.[64] So, either the technique has to be immune to the countermeasures or the countermeasures have to be detected so that assessors know when the results are likely to be distorted.

In my view, the caveats should be taken very seriously. Still, it may be good to look at a quote from Grubin as well regarding polygraph lie detection:

> For the forensic patient, polygraphy offers the opportunity to demonstrate that he is low risk, and it can encourage him to cooperate with treatment and management plans by making it explicit when he is not. It also allows intervention to prevent an increase in risk or relapse in symptoms. Although some may be worried that it will affect the therapeutic relationship with the patient, there is no evidence to suggest such an effect. After all, the aim is to encourage truth-telling rather than to catch the patient out in a lie.[65]

---

61 Megan warns: 'Although one would like to think that free societies could be trusted to use such techniques appropriately, recent events (e.g., the use of torture in interrogations and the increased invasiveness of domestic surveillance by the United States since 9/11) make it clear that such thinking would be naïve'. Meegan, *supra* note 52, at 14.

62 Sartori et al., *supra* note 1; Morse & Roskies, eds., *supra* note 5; Meegan, *supra* note 52; Linden, *supra* note 7; Wolpe et al., *supra* note 50.

63 Haynes, *supra* note 2, at 35.

64 See, eg, Giorgio Ganis et al., *Lying in the Scanner: Covert Countermeasures Disrupt Deception Detection by Functional Magnetic Resonance Imaging*, 55 NEUROIMAGE 312–9 (2011). They reported that 'hemodynamic signals from lateral and medial prefrontal cortices could differentiate deceptive and honest responses but that such differential activation becomes much smaller when participants use a simple covert countermeasure'. Furthermore, they state that, since the 'countermeasures can be learned easily, this study provides evidence that additional research is needed before fMRI-based methods are sufficiently robust to detect concealed knowledge and deception accurately in the real world'. Earlier, Rosenfeld et al. already observed that 'tests of deception detection based on P300 amplitude as a recognition index may be readily defeated with simple countermeasures that can be easily learned'. J. Peter Rosenfeld et al., *Simple, Effective Countermeasures to P300-Based Tests of Detection of Concealed Information*, 41 PSYCHOPHYSIOLOGY 205–19 (2004).

65 Grubin, *supra* note 42.

One of the points expressed in this quote is that the forensic psychiatric patient's words can become more powerful, if the polygraph has shown their truthfulness. In a context in which one's words may be received with some standard skepticism (which is, to some extent at least, true in forensic psychiatric evaluations of defendants), this could indeed be an advantage. Yet, in my view, the possible negative implications for the therapeutic relationship, as discussed, would, among other things, still be an issue as well.

Forensic psychiatric evaluations may regard *past*, *present*, and *future* mental states. The standard evaluation regarding the past concerns legal insanity. The evaluation is retrospective in nature, since it relates to the time of the crime, possibly weeks, months, and sometimes years ago. The standard forensic psychiatric assessment of *current* mental state concerns competency to stand trial: Is the defendant in his current state of mind able to decide about counsel, etcetera? The standard prospective assessment concerns the risk of recidivism. It may well be that the reliability of current mental states in psychiatry is higher than that of future, and particularly past, mental states. So, a general caveat appears to be justified that BMR assessments of past and future should be interpreted with even more caution than those of present mental states. Still, detecting that the person experiences voices here and now (using type IIa BMR) may increase the likelihood that he experienced such voices at the time of the crime, which may be relevant for a judgment about his legal insanity (in particular, since usually no proof beyond a reasonable doubt is required for insanity).[66] In sum, it is likely that the most reliable forensic psychiatric assessments concern present state assessments, while the prospective and—presumably in particular—retrospective assessments will generally be more challenging and uncertain.

There is a further technical issue that has to be considered with respect to forensic psychiatry. Simpson (2008) points to the fact that lie detection results in people with severe mental illness may be different from those in people without mental illness:

> The studies conducted thus far have been carried out on healthy volunteers who were screened for neurological and psychiatric disorders, including substance use. There has been no testing of fMRI liedetection paradigms in juveniles, the elderly, or individuals with Axis I and/or Axis II disorders, such as substance abuse, antisocial personality disorder, mental retardation, head injury, or dementia. It is unclear whether and how such diagnoses would affect the reliability of the approach.[67]

In other words, the presence of a mental illness could *interfere* with the procedure. Clearly, this possibility has to be taken seriously. In practice, this would mean that the BMR technique has to be ecologically validated in people with (severe) mental illness.

Many have argued, for a variety of reasons, that BMR is not ready to be used in court. Yet, as Meegan writes, 'others have noted that neuroscience methods compare favorably, as regard accuracy and validity, to other trial-accepted methods such as psychodiagnostics and clinical psychiatric assessment.'[68] Moreover, Meegan writes, referring to Schauer, 'that even a far from perfect method may be sufficient, in adversarial

---

[66] As Campbell and Eastman (p. 147) state: 'Such neuroimaging evidence may also provide additional information regarding symptoms, which may themselves be causally linked to an alleged offence, or which at least suggest an "offence narrative" contributed to by mental disorder'. Campbell & Eastman, *supra* note 36.

[67] Simpson, *supra* note 48, at 493.

[68] Meegan, *supra* note 52, at 374.

criminal systems, to instil in the jury the reasonable doubt that leads to not guilty verdicts. He recalls that evidence in a criminal trial is asymmetric with the prosecutor required to give convincing evidence to overcome reasonable doubts, whereas the defence may only prove that there is such a reasonable doubt'. This means that even though, on the one hand, the courtroom may set a much higher standard than a research setting (as discussed above), on the other hand, the specific standards of proof in criminal law, may in legal practice diminish the threshold for forensic use of BMR techniques. In addition, Moriarty (2008) has pointed to the fact that, eventually, what counts is not the scientific evaluation of a BMR technique, but the *legal* standard for admissibility of evidence.[69] This is true, yet, we should note, first, that this may not to the same extent be true in all jurisdictions. Second, that there is an intertwinement between the legal and the scientific standard; for instance, in the US legal context *Frye* requires that a certain test is generally accepted in the relevant scientific field. In this way, the law relies, at least in part, on the scientific community to decide about the admissibility of a certain technique.

## 6. CONCLUDING OBSERVATIONS

Forensic psychiatry may well be one of the first areas in which BMR techniques will be used if they become available. In order to structure the exploration of the opportunities and challenges for forensic psychiatry created by BMR, I have used a tripartite theoretical framework of BMR methods. This framework enabled us to recognize and discuss some specific chances and caveats regarding forensic psychiatric evaluations of defendants. Clearly, the application of BMR techniques in forensic psychiatry may have positive effects, basically concerning the reliability and validity of forensic psychiatric evaluations. Meanwhile, three issues are particularly relevant as caveats: countermeasures, confidentiality, and coercion. The first issue, countermeasures, is basically technical: To which extent are such measures possible, and to which extent are they detectable? It refers to the reliability of certain techniques in the forensic setting, where we cannot just assume the subject's honest collaboration. The second issue, confidentiality, is ethical and legal in nature. I emphasized that not only the boundaries of criminal law, but also health law and professional ethics are relevant here for forensic psychiatrists. The third issue, coercion, is both technical and ethical/legal in nature. Regarding the technical aspect of coercive use, not all techniques allow for forced use, because some level of cooperation may be required—but note that they may still be used within the context of a threat or an offer that cannot be refused. Regarding the ethical aspect, medical ethics may well put certain limits on forensic psychiatric coercive use of these techniques. It is essential that all these issues are taken into account, not in the least the medical ethical issues: psychiatrists cannot just rely on the law and legal provisions. They have an independent professional duty regarding the ethics of the decisions they make.

---

[69] Jane Campbell Moriarty, *Flickering Admissibility: Neuroimaging Evidence in the U.S. Courts*, 26 BEHAV. SCI. L. 29–49 (2008). On admissibility of evidence, see also Campbell & Eastman, *supra* note 36. as well as Frederick Schauer, *Neuroscience, Lie-Detection, and the Law: Contrary to the Prevailing View, the Suitability of Brain-Based Lie-Detection for Courtroom or Forensic Use Should Be Determined According to Legal and Not Scientific Standards*, 14 TRENDS COGN. SCI. 101–3 (2010); Katzenbauer & Meynen, *supra* note 13.