



Author's Response to Peer Commentaries: Brain-based mind reading: conceptual clarifications and legal applications

Gerben Meynen*

Department of Criminal Law, Tilburg Law School, Tilburg University, 5000 LE Tilburg, Netherlands
Corresponding author. E-mail: g.meynen@uvt.nl

Glannon and Claydon have written thoughtful peer commentaries on my article about possibilities and perils regarding brain-based mind reading in forensic psychiatry.¹ Based on their comments, I will provide two conceptual clarifications concerning brain-based mind reading, followed by some further thoughts on legal applications.

First, the term 'brain-based mind reading' does not at all imply that the mind is 'based' in the brain.² The term merely expresses that the *reading* is, at least in part, based on brain-derived data.³ In other words, what is brain based is the way in which we *acquire information* that tells something about a person's mental state. In order to further elucidate this point, it may be helpful to look more closely at the notion of mind-reading itself. Hank Greely writes: 'Humans read minds. We constantly try to understand what our fellow humans are thinking and feeling—and how they are going to

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

¹ Walter Glannon, *Some Questions About Brain-Based Mind Reading in Forensic Psychiatry*, J. L. & BIOSCI. (2017); Lisa Claydon, *Brain-Based Mind Reading for Lawyers: Reflecting on Possibilities and Perils*, J. L. & BIOSCI. (2017); Gerben Meynen, *Brain-Based Mind Reading in Forensic Psychiatry: Exploring Possibilities and Perils*, 4 J. L. & BIOSCI. 311–329 (2017).

² The two conceptual clarifications are in response to Glannon's commentary.

³ The notion 'brain-based' lie detection has been used for quite some time now. I consider brain-based mind reading to be the broader term, but conceptually similar.

act'.⁴ This quote brings forward that, first, mind reading is not exceptional, it is what humans normally do, all the time. Secondly, the quote suggests that mind reading is, to some extent, challenging, otherwise Greely would not use the word *try*: apparently, success is not guaranteed. For instance, our boss may say that she much appreciates our work, but we may have difficulty gauging what she really has in mind. Would she be willing to support a promotion?

In addition, it may be good to underline that humans use *all kinds* of information sources to read other people's minds. We may use their words (spoken or written), their facial expressions (including tears), and other aspects of body language. In some cultures, the length and depth of a bow may be taken into account as well. Contextual factors can also be relevant: how long did it take the person to respond to the email, how much time does a person take to talk to us, does the person offer us a drink? The value of each of the sources is not fixed, it depends. For instance, sometimes, facial expressions may be much more informative than the actual words that were uttered. Our mind reading in such a case is to a large extent *facial expression*-based. In other circumstances, a person's bow may be more telling, and our mind reading may be more *bow-based* than *facial expression*-based. Often, the mind-reading process will take into account information from various sources *together*—and we need not be consciously aware of (all of) them. But taking into account a person's bow does not imply that one thinks the mind is 'bow-based', or that one would want to reduce the mind to a bow. Similarly, taking into account information derived from brain tests does not mean that a claim is made about the mind being based in the brain. Finally, if brain-based mind reading would be used, it is likely to be one source of information *among others*, and its value is likely to depend on other sources as well, just like every day mind reading.

The term 'detecting'—as in detecting thoughts or emotions—should also be understood along these lines. Detecting a mental state based on fMRI combined with a computer (machine learning) does not mean that the mental state is 'reduced' to the process of detection. Just like trying to detect or read a person's emotional state by looking at the facial expressions does not mean that the mind would be reduced to the face (or a picture of it, for that matter).

Glannon puts much emphasis on the fact that a correlation between mental states and brain states does not mean that the brain states *cause* or *explain* the mental states. Clearly, I agree, and I never claimed otherwise in my paper. In fact, brain-based mind reading does not presuppose any of this. It is much more practical: it merely aims at contributing to detecting/identifying/reading the correct mental state. Very much like facial expression-based mind reading does not presuppose that the smile 'causes' or 'explains' the person's happiness, or another mental state.

In sum, human mind reading is normal and it makes use of all kinds of information sources. Data derived from a person's brain may, in principle, constitute yet another source. Brain-based mind reading does not mean reading of the 'brain-based mind', but it refers to *mind-reading procedures that rely—to a non-trivial extent—on brain-derived data*. The relevance of brain-based mind reading, in other words, does not in itself rely on the assumption that the mind is brain-based. For the sake of the argument, (even)

⁴ 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 120 (Stephen J. Morse & Adina L. Roskies eds., 2013).

if one assumes that the brain is mind-based, brain-based mind reading can still make sense: the brain could still reveal some information about the mind. And to drive the point home philosophically, being interested in the possibility of brain-based mind reading—including its potential perils for forensic psychiatry—does not necessarily imply (at all) that one is committed to reductionism or reductive materialism. I clarify this point because Glannon seems to suggest otherwise.⁵

The second conceptual issue I would like to address concerns the fact that, *ultimately*, brain-based mind reading is not about identifying ‘neural correlates’ of mental states. Glannon suggests that Marcel Just’s study is about correctly identifying neural correlates of mental states (eg emotions).⁶ A journalist, Meghan Frank, paid a visit to Just’s lab, and she took part in an fMRI mind-reading experiment. Perhaps her account is instructive understanding the actual procedure and its outcome:

Within minutes the computer had analyzed my brain activity compared to other subjects and was ready to guess what I was thinking about. It was given a multiple choice question for each object. Just and Mitchell looked even more nervous than me when the results started to come in. But their smiles grew as did mine when the computer one by one correctly identified what objects I was thinking about in the scanner. I was amazed. I expected the computer to miss a few, but it didn’t. It got them all right.⁷

The first thing to notice is that Meghan Frank doesn’t mention any ‘brain correlates’. She doesn’t write: ‘Within minutes the computer had provided me with all the brain correlates of my thoughts’. Is she mistaken or confused about the outcome of the procedure? Certainly not.⁸ The actual procedure relies on a combination of fMRI brain measurements and a (machine learning) computer that results in a ‘guess’ about a person’s thought or emotion—not in some picture of a ‘brain correlate’.⁹ Frank’s account clearly shows that the role of such a computer in the experiment is crucial: even though the mind-reading procedure is based on fMRI data it is *not limited* to fMRI. So, it is not the case that the fMRI generates images—‘brain correlates of mental states’—and that Marcel Just, looking at the pictures, says: ‘You must have been thinking about a hammer’. In fact, in Just’s experiments, the computer somehow comes up with a result in a way that is not (fully) transparent to us because it concerns machine learning.

⁵ See eg the final sentence of his commentary: ‘More significant is the widely held view that the psychopathology in psychiatric disorders and how it influences behavior emerges from but is not reducible to descriptions of neural activity. If psychiatrists and legal theorists and practitioners continue to hold this view, then future BMR [brain-based mind reading] may have only a limited impact on forensic psychiatry’.

⁶ Glannon writes: ‘As in his analysis of auditory hallucinations, Meynen seems to conflate the neural correlates of emotional impairment and the subjective experience of emotional impairment. He cites a study by Marcel Just and co-authors in which they purportedly “succeeded in identifying emotions based on fMRI scanning” (p. 7). These scans do not identify emotions, only the neural correlates of emotions.’

⁷ Meghan Frank, *Reading My mind*, CBS (2009), <http://www.cbsnews.com/news/reading-my-mind/> (accessed Dec. 27, 2017).

⁸ Even though from a philosophical perspective, some comments might be made about her phrasings.

⁹ Note that Marcel Just is also very much aware of the ‘brain correlates’ issue: one section of his paper is entitled: ‘The Search for Neural Correlates of Emotion’, Karim S. Kassam et al., *Identifying Emotions on the Basis of Neural Activation*, 8 PLoS ONE e66032 (2013). Yet, the title of the paper is very similar to the words I used: ‘Identifying Emotions on the Basis of Neural Activation’. So, being well aware of the ‘neural correlates’ issue is not at all in conflict with conceiving of Just’s technique as a way to ‘identify emotions based on fMRI’ (*and machine learning*).

In the end, as the term already suggests, mind reading—either brain-based or otherwise—is about accurately identifying mental states, not about identifying (brain) correlates of such states. This is not only true for Just's experiment. It is also true for the mind reading we do in everyday life: ultimately, we are not looking for the facial correlates of anger, but, by looking at the person's face we try, for example, to figure out whether our neighbor is still angry about last week's incident. In other words, even though regarding much neuroscience work it is really helpful to point out that it is, ultimately, about brain correlates of mental states,¹⁰ we should realize that the aim, and therefore the design, of (brain-based) mind-reading experiments is different.

In her commentary, Claydon considerably identifies a number of ethical concerns regarding brain-based mind reading for *lawyers*. One of the problems she emphasizes is that 'errors in expert evidence are unlikely to be detectable to the lay person, legal practitioner, or judge'. Consequently, such errors made by experts cannot be corrected (or are very unlikely to be). In my view, this concern for lawyers has, in turn, also ethical consequences for the *expert*. It means that expert testimony should be scrutinized for possible mistakes. If we take this point seriously, it also means that expert testimony should provide a *full* picture because it will not be possible for lay people to fill in the gaps. This is *ethically* relevant to forensic psychiatrists since truth telling is a central ethical principle for the forensic psychiatric expert, as the medical ethicist Paul Appelbaum writes: '*Truth-telling* is the first principle on which the ethics of forensic psychiatry rest'.¹¹ For Appelbaum, this implies that if an expert 'fails to tell the court that his or her conclusions are based on a theory held by only a small minority of peers, or that much evidence exists contradicting the conclusions reached', the expert does not tell the *whole* truth, as he or she should. In addition, Appelbaum writes: 'The psychiatric witness who is being objectively truthful will acknowledge, insofar as possible, the limitations on his or her testimony, including those due to the limits of scientific or professional knowledge, as well as those specific to a particular case (e.g., due to inability to locate records or directly to examine the subject of the evaluation)'. So, in a way, Claydon's point regarding lawyers brings us back to the ethics of forensic psychiatry, in which telling the whole truth is paramount. Of note, Appelbaum adds that staying up to date with the literature is part of being able to tell the truth. Clearly, this can be a challenge regarding neuroscientific developments.

¹⁰ For instance, research in psychiatry may aim at identifying brain correlates of depression or post-traumatic stress disorder. But such a study has a different aim and design compared to an experiment that is about mind reading. To further illustrate the point, ultimately, P300 memory detection experiments do not aim at elucidating the neural correlates of memories, but at (correctly) detecting actual memories (on memory detection using P300 see BRUNO VERSCHUERE ET AL., *MEMORY DETECTION: THEORY AND APPLICATION OF THE CONCEALED INFORMATION TEST* (Cambridge University Press 2011).) To be sure, research could also be designed to identify neurobiological correlates of memories, but that would be a different design and a different research question. Meanwhile, Bennett and Hacker have written a thoughtful book that makes clear that regarding many neuroscience studies it is really helpful to realize that it is about (nothing more than) brain correlates: MAX R. BENNETT & PETER M. S. HACKER, *PHILOSOPHICAL FOUNDATIONS OF NEUROSCIENCE* (Blackwell Publisher 2003). See regarding neurolaw, also MICHAEL S. PARDO & DENNIS PATTERSON, *MINDS, BRAINS, AND LAW. THE CONCEPTUAL FOUNDATIONS OF LAW AND NEUROSCIENCE* (Oxford University Press 2013).

¹¹ Paul S. Appelbaum, *A Theory of Ethics for Forensic Psychiatry*, 25 J. AM. ACAD. PSYCHIATRY LAW. (1997); the other Appelbaum quotes are also from this paper. See on the crucial role of truthfulness in forensic psychiatric testimony, also Douglas Mossman, *Is Expert Psychiatric Testimony Fundamentally Immoral?*, 17 INT. J. L. PSYCHIATRY 347–68 (1994).

Note, meanwhile, that even when brain-based mind reading would start to play a role in legal cases,¹² neuroscience data are likely to form *part of* the information, not all of it. That means that these data could still, in a way, be corrected by other findings in the case. In fact, the role of mind-reading data could be limited even in those cases where they are used. Yet, we should not underestimate the possible impact of ‘supplementary’ data in legal cases.¹³

Claydon draws specific attention to challenges that may lie ahead regarding obtaining consent because it could be difficult to fully explain the ways in which the brain-based mind reading data can and will be used in the future. This is highly relevant since it concerns a *developing* area, and things that are impossible at the moment, may become possible in a couple of years. It is, therefore, impossible to envisage and explain all options for future use—especially if the neurodata would become accessible by different agencies in different countries. In part, such consent problems are similar to those encountered regarding obtaining genetic material¹⁴ (even though, at this moment in time, regarding genetics much more is possible compared to brain-based mind reading). Still, clearly, this would be a point to consider.

Finally, it may be good to ask the question: Why should we think about possibilities and perils of forensic psychiatric use of brain-based mind reading at this point in time? I believe that, basically, the reason has to do with something expressed by Nadelhoffer and Sinnott-Armstrong regarding new neuroscience techniques and possible legal applications: ‘The thoroughly interdisciplinary task of neurolaw is to stay a step ahead of the scientific progress on these fronts so that we can carefully think through the potential implications of introducing new neuroscientific techniques into the courtroom before they arrive’.¹⁵ This expresses the idea behind much neurolaw research also in this journal. Even though brain-based mind reading in forensic psychiatry may sound like neuroscience fiction,¹⁶ time may well have come to start thinking about normative implications of some possibilities and perils.

¹² Note, meanwhile, that my framework of brain-based mind reading (especially type III) is rather broad Meynen, J. L. & BIOSCI. (2017). It also encompasses some diagnostic procedures that are already in use today (for instance regarding brain tumors): diagnosis may reveal something about prognosis, and that may, more precisely, tell us something about a person’s future mental states. For example, the MRI finding of a tumor may tell us that it is very likely that a person’s mental functioning will (further) deteriorate. Therefore, it may contribute to mind reading, at least in a broad understanding of the term. In that respect, it is important to note that neurotechniques have already been used in a diagnostic way in legal cases, in different legal systems (see volume 2, issue 3 of this journal on the use of neurobiological evidence in criminal proceedings in several countries).

¹³ See GERBEN MEYNEN, *LEGAL INSANITY. EXPLORATIONS IN PSYCHIATRY, LAW, AND ETHICS* (David N. Weisstub & Dennis R. Cooley series eds., Springer, 2016).

¹⁴ See eg Frederieke H. van der Baan et al., *Consent in Psychiatric Biobanks for Pharmacogenetic Research*, 16 INT. J. NEUROPSYCHOPHARMACOL. 677–82 (2013). On consent regarding neuroscientific mind reading, see also Adina L. Roskies, *Mind Reading, Lie Detection, and Privacy*, in HANDBOOK OF NEUROETHICS 679–95 (Jens Clausen & Neil Levy eds, 2015).

¹⁵ T. Nadelhoffer & W Sinnott-Armstrong, *Neurolaw and Neuroprediction: Potential Promises and Perils*, 7 PHILOS. COMPASS 631–42 (2012).

¹⁶ Gerben Meynen, *Neurosciencefiction?*, 59 TIJDSCHRIFT VOOR PSYCHIATRIE 454–55 (2017).