HHS Public Access

Author manuscript

AJOB Neurosci. Author manuscript; available in PMC 2020 May 09.

Published in final edited form as:

AJOB Neurosci. 2019 ; 10(1): 32–34. doi:10.1080/21507740.2019.1595785.

Ethical Analysis of "Mind Reading" or "Neurotechnological Thought Apprehension": Keeping Potential Limitations in Mind

Peter Zuk and

Baylor College of Medicine and Rice University

Gabriel Lázaro-Muñoz

Baylor College of Medicine

We appreciate Meynen's examination of ethical implications of using neurotechnologies to decode neural data and make inferences about cognitive processes. Here, we address three issues that we believe limit the relevance of Meynen's analysis. First, at various points, Meynen discusses applications of neurotechnological thought apprehension (NTA) whose technical feasibility is highly uncertain. Second, if some type of NTA technique was reliable and relevant to psychiatry, its use in forensic contexts would likely face significant restrictions in the U.S. and countries with similar legal systems. Finally, what is perhaps the most prominent form of NTA being developed for psychiatry—adaptive deep brain stimulation—is likely to utilize a fundamentally different kind of mental state information than the applications of NTA on which Meynen focuses.

Technological Feasibility and Ethical Analysis

Meynen maintains that neurotechnologies will someday have the capacity to reliably "apprehend mental states, such as: thoughts, intentions, inclinations, biases, desires, feelings, or levels of consciousness" (Meynen 2019, 3). This claim is contestable. While some of the functional magnetic resonance imaging (fMRI) and brain-computer interface results that Meynen discusses are impressive, it is not clear that they license anticipation of the kind of NTA on which Meynen focuses his ethical analysis. In what seems like an attempt to sidestep this reality, Meynen at various points raises the idea that "since this paper is about ethical challenges regarding future NTA use in psychiatry, it is beyond its scope to further discuss in detail the various shortcomings of current techniques" (Meynen 2019, 4). *This is problematic because the ethical analysis of emerging neurotechnologies is highly dependent on the capacities and limitations of these technologies.* It is precisely because the paper is about ethical challenges of emerging neurotechnologies that a discussion of the capacities and limitations of current techniques, and the likelihood of achieving the technological capacities examined is crucial for the analysis.

We value anticipatory neuroethics, but when analyses are not based on the actual or even likely capacities and limitations of emerging neurotechnologies, they risk irrelevance. Discussing Mason and Just's (2016)'s study in which the researchers were able to predict

Corresponding Author: Gabriel Lázaro-Muñoz, PhD, JD, MBE, One Baylor Plaza, Suite 326D, Houston, Texas 77030, glazaro@bcm.edu.

which item from a list of physics concepts subjects were thinking about, Meynen writes that while "it is still far from real-time mind reading, it is suggestive of such a possibility in the near future" despite its multiple-choice experimental setup (Meynen 2019, 3). While Meynen readily admits "that we cannot at present read off a patient's precise thoughts from a scan without having an already determined template for interpreting them" (Meynen 2019, 4), his analysis seems to us to underestimate the significance of this fact. Meynen's envisioned applications of NTA later in his paper fundamentally depart from the multiple-choice paradigm, and this is just one of several potential in-principle problems that may limit the scope and reliability of NTA. This is particularly evident in Meynen's discussion of forensic psychiatry. There, Meynen describes a hypothetical encounter in which "NTA could just register all thoughts that come up in a particular period of time" (Meynen 2019, 8). It simply is not clear that such a technology will ever exist.

That said, Meynen is clearly correct to this extent: these technologies are likely to one day be reliably able to apprehend *some* information about mental states. We now further examine two of Meynen's distinctions, in a way that requires only this more conservative prediction as a working assumption, in order to help inform anticipatory neuroethics on these issues.

Limits on Forensic Use of NTA

Meynen rightly calls attention to the distinction between clinical and forensic applications of NTA. It is important to distinguish these uses, as ethical analyses appropriate to one may be unsuitable for the other. This is a consequence of the different aims of the two fields. Crucially, examinations conducted by forensic psychiatrists do not have the aim of providing treatment, but instead of uncovering potentially legally relevant facts about the subject. However, there is reason to believe that the application of forensic NTA in legal proceedings may be limited, at least in the U.S. and similar legal systems.

Shen (2013), whom Meynen cites, surveys the work of U.S. legal scholars writing on this question, claiming that, "Although there are many doctrinal paths taken, and some notable exceptions, scholars typically find that the Fourth Amendment and the Fifth Amendment both provide protections against mind-reading techniques with neuroimaging" (693). Roskies (2015) raises the intriguing possibility that coercive or compulsory NTA might (also) be precluded by First Amendment considerations. She begins by noting, "Nothing about neuroimaging, which is a measuring technique, interferes with expression, nor does imaging involve any sort of manipulation or restriction of thought content" (Roskies 2015, 686). But, she continues, courts have often held that government action that threatens to cause a "chilling effect" on expression is unconstitutional, and coercive NTA would be likely to have such an effect by threating the very *source* of free expression (Roskies 2015, 686–687).

While it remains to be seen precisely how likely it is that such a chilling effect would occur, compulsory NTA could potentially violate the First Amendment in a more direct way. While Roskies is correct that coercive NTA seems not to directly *restrict* expression, its use might directly constitute *forced* expression. Thoughts, it might be claimed, are a form of

expression: expressions that, in the absence of thought apprehension, are made to oneself. Whether or not private thoughts are indeed expression in this sense, it is plausible that they *become* expression if made non-private by NTA. If that is so, coercive NTA is coercive expression. But precedent in the U.S. recognizes a constitutional right to *refrain* from expression in the civil context: a "right not to speak" (West Virginia State Board of Education v. Barnette [1943], Wooley v. Maynard [1977], Taruschio 2000), in addition to the more familiar Fifth Amendment protection against self-incrimination. For these reasons, it seems unlikely that criminal defendants in the U.S. and similar jurisdictions will be *compelled* to undergo some form of NTA even if the technology does someday reach a level of reliability that courts could consider admissible evidence.

There is however the lingering issue of whether *all* mental states will be treated as on a par. One potential form of mental state apprehension that may prove more technologically feasible than the discernment of complex conscious thoughts is the use of fMRI to identify experiences of *recognition* and *non-recognition* (Rissman et al. 2010). Such a technology could conceivably establish whether an individual recognizes people (e.g. the victim) or things (e.g. a weapon) relevant to a crime, which would make it highly relevant in forensic contexts (Rissman et al. 2010, 9849). One might question whether the having of an experience of recognition (or of novelty) is subject to protection under the First, Fourth, and Fifth amendments.

NTA in Development for the Psychiatric Context

We wish to draw further attention to a kind of NTA information that is fundamentally different, one mentioned, but not discussed in detail by Meynen. In conventional deep brain stimulation (DBS), an implanted pulse generator delivers electrical stimulation to a brain region associated with a particular neurological or psychiatric condition. In next-generation closed-loop or adaptive DBS (aDBS), the goal is to utilize brain activity as biomarkers of pathological emotional states to determine in real time, when and what level of stimulation to apply to regulate these emotional states. aDBS researchers are currently attempting to identify brain signals associated with the symptoms of psychiatric disorders such as obsessive-compulsive disorder (Karas et al. 2019).

Importantly, this application of NTA may also only require discerning *types* of mental state rather than specific *contents* of mental states. In aDBS, what is currently most relevant is that a mental state is of a particular kind: e.g., pathological anxiety. What precisely the anxiety is about or directed toward, if anything, is not clinically *irrelevant*, of course. But it may not be significant for the functioning of the device itself. So long as the device distinguishes pathological instances of anxiety, its functioning need not take into account more particular contents of mental states at all. Thus, while the applications of NTA that Meynen describes are exciting, they may not be part of the current or likely future plans in psychiatry.

The technologies that *are* available or will likely be available in the near future may function in ways that mitigate some of Meynen's ethical worries about NTA. Consider, for example, the issue of confidentiality. While still highly sensitive and deserving of protection, the

information generated and stored by aDBS systems is less likely to be of practical use to overbearing governments and other bad actors than specific contents of thoughts would be. As we have stressed, the technical details of NTA methods can therefore matter a great deal for ethical reflection. Responsiveness to such details remains a crucial guiding principle for neuroethics, including (indeed, perhaps especially) anticipatory neuroethics.

References

Karas, Lee PJS, Jimenez-Shahed, Goodman WK, Viswanathan A, and Sheth SA. 2019 Deep brain stimulation for obsessive compulsive disorder: Evolution of surgical stimulation target parallels changing model of dysfunctional brain circuits. Frontiers in Neuroscience 12: 998. doi:10.3389/fnins.2018.00998. [PubMed: 30670945]

Mason RA, and Just MA. 2016 Neural representations of physics concepts. Psychological Science 27(6): 904–913. doi:10.1177/0956797616641941. [PubMed: 27113732]

Meynen G 2019 Ethical issues to consider before introducing neurotechnological thought apprehension in psychiatry. AJOB Neuroscience [this issue]: 1–20. [PubMed: 31070562]

Rissman J, Greely HT, and Wagner AD. 2010 Detecting individual memories through the neural decoding of memory states and past experiences. Proceedings of the National Academy of Sciences 107 (21): 9849–9854. doi:10.1073/pnas.1001028107.

Roskies AL 2015 Mind reading, lie detection, and privacy In Handbook of Neuroethics, vol. 2, ed. Clausen Jand Levy N, 679–695. Dordrecht: Springer. doi:10.1007/978-94-007-4707-4_123.

Shen FX 2013 Neuroscience, mental privacy, and the law. Harvard Journal of Law and Public Policy 36(2): 653–713. www.harvard-jlpp.com/vols-35-39.

Taruschio AM 2000 The First Amendment, the right not to speak, and the problem of government access statutes. Fordham Urban Law Journal 27(3): 1001–1051. https://ir.lawnet.fordham.edu/ulj/vol27/iss3/10.

West Virginia State Board of Education v. Barnette, 319 U.S. 624 (1943).

Wooley v. Maynard, 430 U.S. 705 (1977).