



Delft University of Technology

## Neurorights Versus Externalism About Mental Content Characterizing the 'Harm' of Neurotechnological Mind Reading

Rainey, Stephen

DOI

[10.1017/S0963180125000106](https://doi.org/10.1017/S0963180125000106)

Publication date

2025

Document Version

Final published version

Published in

Cambridge Quarterly of Healthcare Ethics

### Citation (APA)

Rainey, S. (2025). Neurorights Versus Externalism About Mental Content: Characterizing the 'Harm' of Neurotechnological Mind Reading. *Cambridge Quarterly of Healthcare Ethics*.  
<https://doi.org/10.1017/S0963180125000106>

### Important note

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

### Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

### Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

RESEARCH ARTICLE

# Neurorights versus Externalism about Mental Content: Characterizing the ‘Harm’ of Neurotechnological Mind Reading

Stephen Rainey

Philosophy and Technology, Delft University of Technology, Delft, The Netherlands  
Email: [s.rainey@tudelft.nl](mailto:s.rainey@tudelft.nl)

## Abstract

Neurorights are widely discussed as a means of protecting phenomena like cognitive liberty and freedom of thought. This article is especially interested in example cases where these protections are sought in light of fast-paced developments in neurotechnologies that appear capable of reading the mind in some significant sense. While it is prudent to take care and seek to protect the mind from prying, questions remain over the kinds of claims that prompt concerns over mind reading. The nature of these claims should influence how exactly rights may or may not offer justifiable solutions. Overall, the exploration of neurotechnological mind reading questions here will come in terms of philosophical accounts of mental content and neuroreductionism. The contribution will be to present a contextualization of questions arising from ‘mind-reading’ neurotechnology, and appraisal of if or how neurorights respond to them.

**Keywords:** brain data; externalism; mental content; mind reading; neurodata; neurorights; neurotechnology

## Introduction

Neurorights have quickly become a hot topic of discussion and debate not least in academia and policy circles since the early promotion of the idea in Chile, culminating in a constitutional amendment protecting neural data in 2021.<sup>1</sup> In a 2023 *Atlantic* article, an interview with Rafael Yuste and Jared Genser presented neurorights ideas under the headline “The Right not to have your mind read,” centering the idea of mind reading early in the discourse.<sup>2</sup> At least two senses of ‘mind reading’ are familiar, one from fantasy and science-fiction, the other from psychology. The fantasy or science-fiction idea is probably very familiar—the idea that one’s thoughts might be simply viewed by some arcane mystical means, or recorded somehow by a technological device. This relies on a telepathic notion and a sense of mystery, whereas the psychological sense of mind-reading is more about how people interact, cooperate, and understand one another in general.

In Yuste and Genser’s *Atlantic* article, mind-reading concerns are laid out in terms of mental privacy, and personal identity. Mind reading would affect each of these areas in fairly intuitively obvious ways. Mental privacy, clearly, would be challenged if the mind was somehow open to the view of others. If a room could be peered into by a passer-by it would be hard to consider it a private space, and so too presumably the mind. Personal identity could be impacted through a kind of chilling effect. If one’s mind was in principle open to view, the sense of oneself could be diminished given how one’s hitherto private thoughts might be taken (and potentially out of context). This might be exacerbated owing to the potential cleavage between readable thought and observable behavior too, prompting novel questions about whether brain readings or deliberate behavior ought to be taken as standards for evaluating people.

Neurotechnological mind reading could be seen as a more reliable way to characterize people's thoughts, character, beliefs, and affective states, rather than their testimony.

Fantasy and science fiction might provide inspirational grounds for thought experiments or reflection on deep ideas, though they probably ought not to fixate on us too much. But the notion of mind reading by neurotechnological means is not purely driven by speculation. Ongoing research has provided a body of empirical evidence suggestive of growing mind-reading capacity across a variety of brain recording contexts, seemingly on occasion amounting to a kind of *techno telepathy*. Using functional magnetic resonance imaging (fMRI), Nishimoto and colleagues were able to produce striking reproductions of visual experiences in experimental participants.<sup>3</sup> The results are reproductions of what participants might be seeing 'in their mind's eye.' In other work, using a combination of fMRI and interviews, Horikawa and colleagues were able to identify with some success the content of dreams from sleep research participants.<sup>4</sup> Images and content dreamed about seemingly could be retrieved via the use of brain recording technology. The semantic system of the brain has been mapped also through fMRI recording, providing a means of classifying categories of thought content applicable across different individuals.<sup>5</sup> This suggests at least those types of thought (e.g. whether abstract concepts or concrete objects) could be distinguished by using brain recording, classifying thought without asking a thinker. Combining fMRI with the utility of large language models (LLMs), Tang et al approached the reconstruction of podcast and short animated film clip content from participants, producing text output from thought as represented via fMRI signal.<sup>6</sup> This, on the face of it, looks like the reproduction of a kind of stream of consciousness, retrieved from brain recordings via generative AI.

Each of these apparent examples of mind reading used blood oxygen levels, indicative of brain activity, to infer mental content in one form or another—viewed images, dreamt images, semantic categories of thought objects, and continuous occurrent thought. These clearly touch upon dimensions of mental activity that have privacy and personal identity implications, being relevant to not only the content but also the activity of thought both waking and in sleep. Moreover, these are just four examples from a burgeoning research paradigm from a scientific perspective that itself feeds into private sector interest in mind-reading machines. Between 2010 and 2020 patents in this area tripled.<sup>7</sup>

In apparent neurotechnological mind reading cases, it does not look like the idea of 'mind reading' is utilized in psychology. Mind reading in psychology, according to Goldman,<sup>8</sup> is generally thought of in three competing ways:

- Mind reading as *theorizing* whose animating claim is that "...ordinary people construct, or are endowed with, a naïve psychological theory that guides their assignment of mental states."
- Mind reading as *rationalizing* whose animating claim is that "...the ordinary person is a rationalizer. She assumes that her friends are rational and seeks to map their thoughts and choices by means of this rationality postulate."
- Mind reading as *simulation*, whose main claim is that "...ordinary people fix their targets' mental states by trying to replicate or emulate them... mindreading includes a crucial role for putting oneself in others' shoes."

The decoding of fMRI data does not do theory-theory – no psychological theory is deployed in the decoding. Rather, correlations between mental states and physical brain states are sought. Nor does it look quite like the psychologist's simulation. The LLM approach seems something like simulation, yet it does not replicate or emulate mental states, instead describing apparent thought as per the general architecture of a ChatGPT-style interface. The case of LLM 'mind reading' actually looks most like rationalizing, insofar as it presents apparently cogent connections between associated mental states based on predictions from brain states. But it is not the rationality of a person at stake in this so much as it is the rationality of a domain of language reconstructed from the training corpus (e.g. podcasts). Given this, any quasi-rationalization is disconnected from general human interests, being instead centered on specific genres of language use. The LLM finishes specific stories, based on general stories, given fMRI data as a prompt. In these cases, the technological approach seeks to get a true picture of what mental content is or was occurring at a time. The mind reading at issue is thus more like the sci-fi or fantasy

version of directly peering into a person's mind and in a literal sense reading their thoughts. This technotelepathic sense of mind reading will be taken up and examined here.

The above-mentioned cases provide empirical results that give *prima facie* reasons to think that mind reading by way of neurotechnology is possible and that that possibility may raise novel concerns needing remedies in law and policy. At the base of this are ethical concerns over ideas like those raised by Yuste and Genser in the Atlantic over privacy and personal identity. But there is also good reason to examine this jump to 'mind reading' as the most apt concept at stake here. Do the four examples outlined, and other similar cases, amount to mind reading or to something else? This is important because to get a handle on the right remedies for problems, we really ought to get the problems straight first. To hone in on the issue it is worth raising a salient and fundamental question: Is the brain a sufficient basis for identifying mental content? If this question can be answered in the affirmative then neurotechnology could present mind-reading questions in need of responses. If not, then we need to reconsider the framing of concerns about neurotechnology and the mind, which perhaps requires different responses. In what follows, I suggest the answer to the questions is a straight, or a qualified, 'no.' As such, 'mind reading' concerns are best reconsidered in other terms. First, it will be important to show that mental contents are not purely resident within the confines of a skull and that at least some form of externalism is required to properly characterize mental content.

### Externalism about mental content

This section is not intended to provide full arguments for each of the positions put forward, but rather to highlight some key points made by Hilary Putnam, Tyler Burge, Andy Clarke, and David Chalmers, as well as considerations from an enactivist position. This will show that mental contents' confinement within the head is not as secure as might be presumed. If mental contents are essentially world-involving somehow, then the brain alone is not sufficient to identify mental contents. So, if an externalist view of mental content is better than an internalist one, mind-reading problems either aren't real or ought to be recharacterized in more apt terms. In fact, the aim here is not to establish externalism once and for all, which would be tricky in a short article. Rather, the aim here is to produce enough doubt over neuroreductionism and the thought that externalism can be seen to have at least some truth to it. With this weaker aim established, it will be sufficient to show claims that mind reading ought to be rephrased for specificity's sake. This will also prompt reflection on how to deal with purported mind-reading questions, such as those over privacy and personal identity. First, a brief outline of the four dimensions of externalism:

#### Natural

In the 1970s, Hilary Putnam developed so-called Twin Earth thought experiments to demonstrate the need for semantic externalism, that is, to show meanings of words cannot be given definitively by means of psychological explanation.<sup>9</sup> Meanings, in Putnam's account, are world-involving. To summarize this view imagine a 'Twin Earth,' just like the Earth we know in almost every respect. However, Earth Gold and Twin Earth Gold have different chemical structures. A person suddenly transported from Earth to Twin Earth might not change any of their beliefs about the ring they wear on their finger, but they might now be mistaken about their wider beliefs. "My gold ring is made of that same material" might be true on Earth when comparing one's ring with one of the contents of Fort Knox. On Twin Earth, and the contents of Twin Fort Knox, this is false on account of the different structures. 'Gold' means gold straightforwardly on Earth, but something different on Twin Earth. If someone swaps my familiar ring for its Twin Earth counterpart the ring I now wear is not gold at all, regardless of the word and my belief about it, because Twin Earth gold is not the same thing as Earth gold. I am susceptible to being mistaken in my meanings as well as my beliefs owing to the nature of the world around me. Some factors beyond the content of the mind are required to individuate some of my thoughts and some truths. Putnam summarizes this view as follows: "Cut the pie any way you like, "meanings" just ain't in the head!"<sup>10</sup>

### Social

Beyond the natural environment and the causal connections between it and thought of it, various beliefs appear enmeshed among the linguistic and conceptual resources of language and are made possible only in that web. This sociolinguistic aspect of mental content means that different mental content is realizable in different minds, depending on social inheritance and linguistic competence. The widely referred to Sapir Whorf hypothesis casts the concept of linguistic relativity as influential upon thought and cognition in general to different degrees.<sup>11</sup> Edward Sapir and Benjamin Lee Whorf differed in their emphasis upon how deterministic their hypothesis was, but overall the contention is that language does not just reflect reality, as might pretheoretically be assumed, but shapes the experience of it. It can be pursued further that language also shapes thought itself, bringing different constraints and scaffolds for different ideas, depending on the conceptual contents available in language. If true, this would place some of the semantic differences between linguistically identical terms outside the person using the language.

Following an example from Tyler Burge, it even seems that given the socio-cultural dimensions of words and their meanings, each of us can be wrong about what we believe (and hope, doubt, wish for, etc.).<sup>12</sup> Let us imagine Donald is told he has gout. Donald believes that arthritis is a kind of joint inflammation, but not the same as the inflammation brought on by gout. Donald thus believes he has gout but does not believe he has arthritis. If Donald knew better the concept of arthritis, he would know that it is an inflammation of the joints, and so he would know he did in fact have arthritis. Donald's loose grasp of the concept of 'arthritis' means he is mistaken in his belief that he does not have it, having instead gout. But on Burge's account, we can imagine a world in which gout is not part of the concept of 'arthritis.' In this other context, arthritis might be any inflammation of the joints, *except gout*. Donald in this world is correct in his belief that he has gout and not arthritis. Between these two worlds, there are no physical changes in the state of Donald, yet there is a truth-value change regarding his beliefs. If Donald were to be shipped from the first to the second context, his belief would become true but nothing else in the world would change, *and he would not know it*. In a palpable sense, Donald would not know the content of his own beliefs. Without knowing which world Donald was in, there would be no reflective way to know whether his thoughts about arthritis were true or false. Donald's mental content, and mental content in general, depends on external factors over and above any reflection he might make upon his own states or beliefs.<sup>13</sup>

### Active and enactive

Clark and Chalmers'<sup>14</sup> extended mind hypothesis is another means of understanding mental content as going beyond the confines of the skull, such that a technotelepathy approach is insufficient for 'mind reading'. Clark and Chalmers argue for the concept of "active externalism," whereby external objects play an active role in cognitive processes. If a person can substitute a notepad or the contents of a hard drive for their memory, and they come to rely on and trust that substitute as they would the cognitive process, then active externalism suggests the system including the world is equivalent to the cognitive process. Memory can be described in terms of private reflection, but also as the use of an external system of objects or devices. This being so, if mental content is thought to be actively distributed among items in the world, notepads, hard drives, etc., then the role of the brain itself is at least partly dealing with processing information rather than representing content. Reading the activity of the brain, in this case, is not enough to determine content, since the brain is just one part of a system extending into the world at large.

Relatedly, especially in terms of recognizing the mind as world involving, we might follow enactivists<sup>15</sup> in thinking of mental content like 'the experience of seeing a pink square' as a learned activity. We have plenty of empirical evidence to suggest our eyes do not take in the pink square as such. Our visual experience is knitted together from eyes darting in saccades, edges observed, colors filled in from memory and expectation as much as the irradiation of the retina from photons of light. Anil Seth goes as far as saying that experience is a 'controlled hallucination'<sup>16</sup> constructed by our brains. From an enactivist perspective, the control in this suggestive image from Seth is a skilled enaction of perception

and cognition. Certainly, optical illusions reveal sight as not a straightforward observation of objective reality as they exploit what we know of our physiological limitations and cognitive ingenuity. So when we do see a pink square, we are really enjoying the learned activity of accounting physically for the kinds of movements we need to make to produce a steady observation, while cognitively we are filling in a lot of blanks in order to perceive shape and color.

The pink square ‘in our minds’ is an achievement hanging on a lot of skilled experience, based on a sensorimotor dynamic such that perception is something we do, not something we undergo.<sup>17</sup> That makes mental content again both dependent on the world, and on having adopted various behavioral routines necessary to satisfactorily register experience at all. Alva Noë argues that “...for at least some experiences, the physical substrate of the experience may cross boundaries, implicating neural, bodily, and environmental features.” This means that, just like Clark and Chalmers’ view of cognitive processes potentially inhabiting the world beyond the brain, experience more generally may involve the wider world. The brain itself is not enough to determine the content of this scheme either, since the world, the body, and learning skills are all required to have any given experience. No act of brain reading will provide a complete story concerning mental content, or the content of experience on these views.

In summary, there are compelling analyses of perception and experience that indicate a central role for the world at large, sociocultural contingency, and learned skill in the production of mental content. These analyses, if any part of them is at all true, provide sufficient grounds to doubt that the reading data from the brain is sufficient for reading content from the mind:

- **Natural**
  - Mental contents, e.g. meanings of words, rely on causal interactions with the natural world
- **Social**
  - Beliefs and language interact, constraining or conditioning worldview and cognition, as a product of sociocultural and historical contingency
- **Active and enactive**
  - Vehicles of conscious experience include more than what’s *in the head*
  - Mental contents are skilled interactions between body and environment – a ‘sensorimotor dynamic’

Decoding images from the brain then relies not just on neural activity but also the environment which the person has occupied – for the Twin Earth residents on Earth looking at or dreaming about gold a determination that their mental content is ‘gold’ is false. Less abstractly, though, the same holds: a causal connection between the image seen or dreamt of has to exist. In a ‘proper’ way to get to content. That connection must be established in addition to any image even when we assume the accuracy of the technology. Aside from the natural world, there are also commitments made about sociolinguistic inheritance when seeking to determine mental content. Importantly, and especially when socio-cultural shifts are at stake, mental content might not be transparent to the knower themselves, let alone the interpreter of brain data who might seek to ascribe occurrent thought or infer belief. Without reference to or knowledge of sociolinguistic inheritance, this could at the very least be a tall order. And even then, the brain may not be the stage on which the mental show supposedly plays out anyway, with the world at large being implicated in perception, cognition, and wider experience. This introduces the complication of judgment in characterizing the data recorded from the brain.

Cases of purported mind reading aren’t like cases of reading words from a page as mental content is not there like the words are on the page. Even though interpretation is required when reading words from a page, the words themselves are present to view. But with a complicated story about mental content, it is not clear the content is ‘there’ to be read.<sup>18</sup> The experience of the target, the classification of the machine, the conceptual and theoretical presuppositions latent in its code and algorithms, the interpretive judgment of those reading the results of the machine, and so on, each contribute to what image is produced. And these may easily fail to map onto one another neatly. Mind reading efforts depend on much more than the specific activity of neuronal circuits and knowledge of experimental conditions.



### *Prediction and protection of data*

The foregoing discussions are intended to raise doubts over the possibility of technotelepathy by means of neurotechnology. They have aimed to do this by showing the brain is not a sufficient basis for identifying mental content. Nevertheless, it is clear from experiments like those already mentioned that clever interpretation of brain activity nevertheless permits predictions about mental content. It remains now to reframe the issue of mind reading as 'mental state prediction' based on neurotechnology.

Mental states are predicted from brain states in experimental conditions through the association of brain activity with observed perception, cognition, affect, and so on. As such, it is more like data science than telepathy. We already know a lot about the possibilities and dangers of data science when it comes to making predictions about people. Using big data can help to model groups, predict behavior, preferences, health outcomes, and so on, to help streamline policy investments in healthcare.<sup>19</sup> But we also know of the risks and there are established discourses on how data can contribute to and exacerbate social ills through too zealous regard for data.<sup>20</sup> There are choices made when constructing data sets and using them that can be overlooked in thinking of data in general as somehow objective.<sup>21</sup> In the emerging case of mental state prediction, it will be important to keep the progress made in scrutinizing data to the fore. Applications of mental state prediction could be informative from a human research perspective in empirically exploring mind/brain connections. This could bring benefits in psychiatric contexts, by providing new tools for exploring mental illness. There could also be communication applications or more farfetched collaborative and collective thought explorations. But in each case, the basis would be prediction, not reading, of mental state. This distinction and the basis in data must be emphasized in any conversation that veers into 'mind reading' territory. This is not just for the sake of accuracy, but also because it alters the stakes, the concerns, and the remedies appropriate to them.

In looking for data to identify mental states, a step is made away from the more usual modes of mind reading associated with psychology, as noted above. Centrally, a move to data in this context could mark a move away from asking people about their mental states, thereby minimizing their subjectivity as having a role. Across a spectrum of philosophical positions, in no particular order, the importance of one's subjectivity, or first-person perspective, can be characterized in terms of: what it is like to be a specific entity;<sup>22</sup> the way individuals conceive of the world *in particular*;<sup>23</sup> the 'I think' that accompanies all one's experience;<sup>24</sup> the specific perspective from which experience is had;<sup>25</sup> the locus of judgments, decisions, responsibilities one makes or takes on;<sup>26</sup> one's positionality within and shaping through socio-cultural norms.<sup>27</sup> These are each way of considering how an individual lives their own experience, including their mental state. The intrusiveness of technotelepathy could be seen as providing reasons for new laws and rights such that no technology could legitimately be put to use to read minds. The stakes here would indeed include mental privacy and personal identity. Mental prediction might well be seen as intrusive too, but the main concern should be seen in terms of *misinterpretation and mischaracterization*, with stakes in terms of subjectivity.

The specific harm of predicting mental states or content on the basis of data is the reduction of persons to data streams or the predictions of a model. This reduction comes at the cost of their subjectivity. Overreliance on, or the misinterpretation of, data mischaracterizes subjects and overlooks the subjectivity that is so important to individuals. This thereby serves to eclipse what it is like to be a specific person, the way that person conceives of the world in particular, their specific perspective, their status as a locus of judgments, their orientation toward socio-cultural norms, and so on. Overlooking these facets of importance regarding mental content and favoring predictive data analytics is the basis for the harms risked in 'mind reading' contexts.

Especially difficult questions concerning the ethics of mind-reading technologies arise once their application is envisioned beyond the lab and its already highly regulated context of human research. In private enterprise, medical, governmental, or law enforcement contexts, for example, there are potential threats to mental privacy and personal identity in the ways suggested by Yuste and Genser. But given the threats aren't realized via technotelepathic means. They should be seen as being generated from the nature of mental prediction based on processing brain data, and the potentially reductive applications to which it might be put. This is a different issue to that of techotelepathic mind reading, but one no less

problematic. Importantly, the difference between the two varieties of concern should prompt different responses. Mental state prediction, being based on data, can and ought to be controlled via data regulation and close scrutiny of applications that include subjectivity-reducing risks.

## Conclusion

In the context of neurotechnology development and mind reading mental privacy is, in a strict sense, not newly infringed upon. This is because technotelepathy is untenable. One's personal identity and sense of self are protected best from neurotechnological activity by way of protecting data recorded from brains – neurodata protection. This kind of protection, when effective, inhibits the extent to which predictions can be based on brain data and so cuts off the step from brain state prediction to mental state prediction. The reduction of persons to neurodata streams challenges subjectivity unfairly, and controlling neurodata effectively mitigates those challenges. Novel neurorights do not appear to be necessary in this context since no radically novel threats to the mind are realized by neurotechnological mind reading. The discourse around neurorights is valuable in rightly identifying important areas for attention as this technology evolves. It has not to date foregrounded correctly the specific harm from 'mind reading' that neurotechnology brings, which is the reduction of persons to data streams to the detriment of their subjectivity. Neurorights could, of course, protect people from this harm. But if their protections against the reading of the mind were based on an erroneous characterization of the threat, they would not be adequately justified.

## Notes

1. Neurorights Initiative. *Chilean Senate approves regulation of NeuroRights*; 2020; available at <https://nri.ntc.columbia.edu/news/unanimously-chilean-senate-approves-regulation-neurorights> (last accessed 25 Mar 2021).
2. Andersen R. *The Right to Not Have Your Mind Read*; 2023; available at <https://www.theatlantic.com/technology/archive/2023/08/mind-reading-brain-data-interrogation-mri-machines/675059/> (last accessed 25 Nov 2024).
3. Nishimoto S, Vu AT, Naselaris T, Benjamini Y, Yu B, Gallant JL. Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology* 2011;**21**(19):1641–46.
4. Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y. Neural decoding of visual imagery during sleep. *Science* 2013;**340**(6132):639–42.
5. Huth AG, De Heer WA, Griffiths TL, Theunissen FE, Gallant JL. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 2016;**532**(7600):453–58.
6. Tang J, LeBel A, Jain S, Huth AG. Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience* 2023;**26**(5):858–66.
7. Bertoni E, Ienca M. *The privacy and data protection implication of the use of neurotechnology and neural data from the perspective of Convention 108* 2024.
8. Goldman AI. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford; New York: Oxford University Press; 2006, at 4.
9. Putnam H. The meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science* 1975;**7**:131–193.
10. see Note, 9, Putnam 1975, at 144.
11. Hussein BA-S. The sapir-whorf hypothesis today. *Theory and Practice in Language Studies* 2012;**2**(3):642–46.
12. Burge T. Individualism and the mental. *Midwest Studies in Philosophy* 1979;**4**:73–121.
13. There is much more to this, e.g Davidson D. Knowing one's own mind. *The Twin Earth Chronicles*. Routledge; 2016:323–41 for deeper consideration.
14. Clark A, Chalmers D. The extended mind. *Analysis* 1998;**58**(1):7–19.
15. Ward D, Silverman D, Villalobos M. Introduction: The varieties of enactivism. *Topoi* 2017;**36**(3):365–75.



16. Seth AK. Our inner universes. *Scientific American* 2019;**321**(3):40–47.
17. Noë A. *Action in Perception*. Massachusetts Institute of Technology; 2004.
18. Davidson pursues this point nicely in analysing accounts of subjectivity and first-person authority, but it is a bit out of scope here. See Davidson D. Knowing one's own mind. In: *The Twin Earth Chronicles*. Routledge; 2016:323–341.
19. Marino S, Xu J, Zhao Y, Zhou N, Zhou Y, Dinov ID. Controlled feature selection and compressive big data analytics: Applications to biomedical and health studies. *PLOS ONE* 2018;**13**(8):e0202674; Craglia M, Hradec J, Troussard X. *The Big Data and Artificial Intelligence: Opportunities and Challenges to Modernise the Policy Cycle*, *Science for Policy Handbook*. Elsevier; 2020:96–103.
20. Zuboff S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Paperback edition ed. Profile Books: London; 2019; O'Neill C. *Weapons of Math Destruction, How Big Data Increases Inequality and Threatens Democracy*. Penguin; 2016.
21. Bollier D, Firestone CM, *The Promise and Peril of Big Data*. Aspen Institute, Communications and Society Program; 2010.
22. Nagel T. *The View From Nowhere*. New York: Oxford University Press; 1989.
23. Clark A. Spreading the joy? Why the machinery of consciousness is (probably) still in the head. *Mind* 2009;**118**(472):963–93.
24. Kant I, Guyer P, Wood AW, Kant I. *Critique of Pure Reason*. Cambridge; New York: Cambridge University Press; 1998.
25. Williams BAO. Personal identity and individuation. *Proceedings of the Aristotelian Society* 1956;**57**:229–52.
26. Moran R (ed.), Précis of 'authority and estrangement: An essay on self-knowledge'. *Philosophy and Phenomenological Research* 2004;**69**(2):423–26.
27. Heal J. Moran's 'authority and estrangement'. *Philosophy and Phenomenological Research* 2004;**69**(2):427–32.