ORIGINAL PAPER

Neurotechnological Applications and the Protection of Mental Privacy: An Assessment of Risks

Pablo López-Silva · **Abel Wajnerman-Paz** ·
Fruzsina Molnar-Gabor

Abstract The concept of mental privacy can be defined as the principle that subjects should have control over the access to their own neural data and to the information about the mental processes and states that can be obtained by analyzing it. Our aim is to contribute to the current debate on mental privacy by identifying the main positions, articulating key assumptions and addressing central arguments. First, we map the different positions found in current literature. We distinguish between those who dismiss concerns about mental privacy and those who endorse them. In this latter group, we establish a further disagreement between conservative and liberal strategies to protect mental privacy. Then, we address the first discussion by articulating and criticizing different skeptical views on mental privacy. Finally, we try to identify what are the unique features of neural data and examine how they may be connected to the ways in which neurotechnological mindreading could put mental privacy at risk. We suggest that even if neural data is unique, it may not require new strategies to protect people from its misuse. However, identifying the special features and risks of neurotechnological mind-reading is necessary for the second discussion on mental privacy to properly take off.

P. López-Silva
School of Psychology, Faculty of Social Sciences, Universidad de Valparaíso, Valparaíso, Chile

P. López-Silva
Millennium Institute for Research in Depression and Personality, Valparaíso, Chile

A. Wajnerman-Paz (✉)
Instituto de Éticas Aplicadas, Pontificia Universidad Católica de Chile, Santiago, Chile
e-mail: abelwajnerman@gmail.com; abel.wajnerman@uc.cl

A. Wajnerman-Paz
Neuroética Buenos Aires, Buenos Aires, Argentina

F. Molnar-Gabor
University of Heidelberg, Heidelberg, Germany

Neural Data and Mental Privacy

The search for medical neurotechnological applications has led to the creation of novel ways to register and record neural activity [1–5]). The development of techniques such as Deep Brain Stimulation (DBS), Transcranial Magnetic Stimulation (TMS), and Brain-Computer Interfaces (BCIs) are opening new possibilities of treatment to conditions such as Parkinson's disease, Brain Strokes, and Paralysis among many others [6–10]. Over the last years, a number of ethical worries have been raised due to the fast development of neurotechnological applications in the private sector and how these applications could impact different dimensions of society such as,

for example, education, labor, entertainment, military systems, data management, among many others [11–16]. In this context, part of the scientific community has started to worry about the ways in which *mental privacy* could be at risk (e.g. [17–19]). The notion of *mental privacy* refers to the degree of control that subjects should have over the access to their own neural data (or, at least, that informed consent is necessary for obtaining access) and to the information about their mental processes and states that can be obtained by analyzing such data [17, 19–24].

Two fundamental disagreements seem to characterize the current state of the discussion about mental privacy; the first one has to do with whether mental privacy is really at risk. In this context, a first group claims that—for philosophical and empirical reasons—mental privacy is not really at risk, dismissing the very worry about how to protect it (e.g., [25–27]). Let's call this the *skeptic view*. A second group will agree that neurotechnological applications pose real threats to mental privacy (see [15, 20, 23]). The second disagreement concerns this latter group. Agreeing that mental privacy should be protected by legal means is one thing but agreeing on how to protect it is quite another. This debate also seems to remain open in current literature [19, 20, 28, 29]. For some, citizens would be already protected from the type of threat raised by neurotechnological applications by current laws, or by simply adding minor specifications to currently existing legal frameworks (e.g., [30–35]). Let's call this the *conservative view*. For others, the type of threats to mental privacy posited by commercial neurotechnological applications might not be covered by current legal frameworks or international treaties, and for this reason, new regulatory frameworks should be developed (e.g., [17, 20, 24, 28, 33, 36–38]). Let's call this the *liberal view*.

Each view in these debates faces different challenges. Arguing that mental privacy is not in need of protection requires the skeptics to show that neurotechnological developments either do not provide access to (relevant) neural and mental processes or that the access they provide is not a real threat to users. In other words, the skeptic needs to counter a very plausible and intuitive worry. In turn, arguing that mental privacy needs protection, but new regulatory frameworks are not needed, requires the conservative to clarify specifically how existing laws are able to mitigate the neurotechnological risks that did

not exist when they were drafted. In addition, they also need to deal with the question about how such laws would specifically operate in society throughout specific legal systems (typification of related crimes, establishment of aggravating circumstances, degrees of severity of infringements, what constitutes evidence, among many others). Finally, the liberal needs to connect specific neurotechnological risks to specific gaps in current regulatory frameworks and explain how those gaps need to be filled (adaptation of currently existing frameworks, creation of new human rights, establishment of analogies with currently existing laws, among others options). As it stands, the debate about mental privacy protection seems to be far from settled.

After contextualizing the general discussion about mental privacy (Section "Contextualizing the Debate: Neurotechnologies and the Need for Neurorights"), we oppose the skeptic view by replying to its main philosophical and empirical arguments (Section "Skepticism toward Mental Privacy Threats"). Then, we propose that a way of disentangling the debate between liberals and conservatives implies stepping back and determining whether there are risks that are unique to neurotechnological developments (Section "What is Unique About Neurotechnological Mind-Reading?"). If neurotechnological risks are identical to those associated with other technologies that also provide access to mental processes and that are already adequately regulated, then the discussion will not take off and the conservative view can be endorsed without much discussion. In contrast, if there are risks that are uniquely associated with neurotechnological mind-reading, then we need to have a substantive discussion regarding whether these new risks can be adequately addressed by current regulations, which were not specifically developed to fulfill this task. We approach this issue by critically examining the uniqueness of the type of data on which neurotechnological mind-reading relies. We then argue that these unique features introduce new risks that are not associated with other forms of technological mind-reading. We finish by claiming that identifying what is unique about neural data is only the first step in the discussion between liberals and conservatives (Section "Concluding Remarks: Should Unique Threats Induce The Creation of Neurorights?"). In order to illustrate this, we examine the liberal proposal championed by the *MorningSide Group*,

then known as the "Neurotechnology Task-force"– NET henceforth –, that proposes that some of its unique features make neural data analogous to *organic tissue*. As a consequence of this, the proposal suggests that neural data should be legally treated as organic tissue [28, 29]. We argue that the features that the NET's proposal appeal to cannot ground that specific legal analogy.

Contextualizing the Debate: Neurotechnologies and the Need for Neurorights

The worry about the protection of mental privacy is part of a broader concern about the different ways in which subjects could be affected by the misuse of neurotechnologies in medical and non-medical contexts. In 2014, former US President Barack Obama warned the international community about the impact that neurotechnological intromissions in the brain's dynamics could have on data privacy, agency, and personal responsibility [39]. The main priority of large-scale neurotechnology projects has been the search for novel medical treatments. However, as [40], p. 11) suggests: 'it would be naïve to think that companies that have invested considerable amounts of money on these projects might not want to find specific ways to secure the return on that investment by creating commercial uses of those technologies'. Here the idea is not that neurotechnological applications should be rejected per se, but rather that their wide range of potential applications might give rise to a number of unforeseen practical and ethical problems.

Motivated by this issue, interdisciplinary communities have started to discuss different regulatory strategies through which interests such as agency, privacy, and liberty might be protected from misuses and abuses of invasive and non-invasive neurotechnological devices [17, 24, 37, 38, 41–45]. In this context, the NET—formerly known as Columbia University's *MorningSide Group*—suggested that neurotechnological developments are revolutionizing the way in which neurological conditions are treated, however, at the same time: '[neuro]technology could also exacerbate social inequalities and offer corporations, hackers, governments or anyone else new ways to exploit and manipulate people' [38], p. 160). One of the main problems identified by the NET is that, apparently, existing legal frameworks would be insufficient to

deal with this type of threat [28, 29, 38], and for this reason, the group proposed the concept of *neuroright* as a required legal resource that should accompany neurotechnological developments in order to protect 'people's privacy, identity, agency, and equality' [38] p. 159).

The NET has developed one of the most widely discussed proposals about the creation of specific rights associated with neurotechnological development, suggesting the creation of fundamental neurorights based on four concerns: (i) The right to *privacy & consent* refers to the right for a person to keep their neural data private, ensuring that the ability to opt out of sharing is the default mode of ownership of such data. (ii) The right to *agency & identity* is the right for a person to preserve their sense of agency and human identity in light of potential changes produced by the use of neurotechnologies. (iii) The third concern has not been crystalized in the creation of a specific right yet. This concern refers to the possibility of unequal *augmentation* of cognitive and physical functions through the use of neurotechnologies and how this augmentation would produce new types of inequalities. (iv) The last concern refers to the way in which certain *biases* could become embedded into neurotechnological devices and how this might be avoided. Although the creation of the concept of neurorights is a step forward when it comes to discussing ways to protect subjects and societies from potential misuses of neurotechnologies, the ways in which such worries can be integrated into specific legal frameworks are still under debate (see for example, [15, 20, 31, 46, 47]).

Lately, the neuroright crusade pioneered by the NET has been widely criticized. For some people, the proposal is replete with highly speculative predictions about the way in which neural data might be used and collected, thus decreasing the general appeal of the whole debate (see, for example, [31]). In the same vein, it has been claimed that the potential inclusion of terms with a deep philosophical baggage—such as agency and free will—into the law leads to a number of practical difficulties that make the regulatory task nearly impossible [30, 48]. For others, the creation of specific neurorights is redundant, as current frameworks could be extended to the relevant cases without neglecting their specificity [35]. The idea is that, in the same way in which new forms of murdering a person do not require new ways of protecting the right

to life, new ways of trespassing privacy would not require the creation of new laws. All forms of privacy are already protected by the law. From this point of view, the NET proposal could trigger an 'inflation' of rights (see [31, 49]). In this vein, for some authors the right to mental privacy could be integrated into the already existing frameworks for the protection of the right to privacy, freedom of thought, and freedom of expression [33]. Similarly, Bublitz [31] claims that 'insofar as some forms of neurodata are not covered but should be so, one may insert "neurodata" to Article 9 [of the European Union`s General Data Protection Regulation—GDPR], next to other types of data such as genetic data. No need for further reforms' (p. 7). Even if this is true, it is an issue that needs to be discussed. It is possible that the current protection provided to sensitive data based on the applicable data protection laws is insufficient to prevent potential threats related to neural data misuse in a proportionate manner, and this possibility cannot be dismissed until the risk assessment of the use of neurotechnologies is well established.

Skepticism toward Mental Privacy Threats

Not everyone agrees that current neurotechnological applications pose relevant threats to mental privacy. If this is not the case, then no discussion about how to protect mental privacy is justified. Arguments for this skeptic position have been presented in both the philosophical and empirical domains. In what follows, we examine and reply to some of these arguments, in order to motivate the idea that worries about mental privacy infringements are well justified and need to be discussed.

The Philosophical Domain

Until very recently, the mind was regarded as our most inner and private space for thought, the last bastion of our freedom and privacy. However, the unique features of the data about brain activity and structure—neural data—make possible for certain neurotechnologies to undermine the mind's opacity, its inaccessibility to others, more than any other existing mind-reading technologies and techniques. This opacity is what may give us control over which beliefs, values, feelings, etc. we share publicly; this opacity

makes it possible to selectively exteriorize our inner lives, and therefore, to exert privacy as a mental capacity (Altman 1976, [22, 50]). However, different arguments have been proposed to challenge the idea that neurotechnological mind-reading can undermine mental privacy in a relevant sense. Here we will mention two main philosophical arguments which push in opposite directions.

According to a first argument, mental privacy cannot be undermined because the kind of mental states we would like to keep private *are private by definition* [25]. The main idea is that the key mental states that are at the center of the worry about mental privacy—such as our personal experiences, feelings and thoughts—are essentially subjective, and therefore, they can only be first-personally accessed. This privileged mode of epistemic access is also associated with the idea that there would be 'something that is like' to be in those states only for the subject undergoing them [51–53], this establishing an ontological and epistemological boundary between the 1rst and the 3rd person. So, although we can access the behavioral and neural correlates of conscious mental states and grasp their main psychological features (3erd person access), an external observer has no direct access to those subjective states themselves [25]. As a consequence, the kinds of mental privacy infringements that matter to us are in principle impossible.

A second line of argument establishes that it is not the case that mental privacy infringements are impossible, but rather that they are too common to be ethically problematic. The idea is that we are all endowed with a natural capacity for mind-reading, usually known as the "theory of mind" [54, 55]. This capacity is defined as the psychological ability to access other people's mental states (intentions, beliefs, desires) on the basis of their overt behavior, gestures, expressions, etc. We exert this ability almost all the time, as it is the foundation for our ability to socially interact and predict behavior. However, we do not regard this daily access to other people's minds as ethically problematic and most often we do not ask for consent (nor expect others to ask for our consent) before using this capacity. Thus, if we do not regard natural mind-reading as ethically problematic, it seems reasonable to conclude that technological mind-reading is not problematic either [27]. Although these arguments might have some traction, there are a number of ways to counter them.

Regarding the first argument, even if it is true that first-person conscious mental states are epistemically private, there is a close relationship between them and their biological underpinnings. In fact, it is plausible to say that without their specific biological conditions, such states would not even exist in humans the way they do. Secondly, conscious mental states cannot be reduced to their private subjective character; they are an assemblage of phenomenal and physical properties, even if the ontological relation between them is still unclear. In this sense, brain activity or behavioral expression is not less part of a conscious mental state than its phenomenal character. The issue here is that neurotechnological devices could allow access to one of the most fundamental portions of our inner mental life i.e. its neurofunctional footprints. So, even if it is true that we cannot directly access someone else's *qualitative subjective character of experience,* the very possibility of access to the way that the brain behaves while undergoing certain specific first-personally accessible mental states offers the possibility of intromissions into one of the necessary conditions for a subject's private realm (*see* [56]*,* for example).

Even if neurotechnological mind-reading cannot access the *form* in which certain contents appear in a subject's field of awareness, neurotechnological devices might still be able to access fundamental aspects of the *representational contents* of those states via mind-reading; this would entail trespassing the subject's exert of her ability to privacy as a mental capacity. Take the case of recently developed decoding methods that define a semantic space in the cerebral cortex that represents thousands of different mental categories as a result of the many possible combinations of neural activity patterns recorded by fMRI (e.g., [57–59]). These methods have been shown to be able to identify, from a large set of completely new natural images, which particular image a person has seen. In other words, such methods have been able to access private representational content even if they are not able to directly access what it is like to experience that content in the experimental subject's field of awareness. More recently, it has been possible to decode the categories of objects perceived by a person while watching a movie or dynamic image (e.g., [60, 61]). So, even if it is true that we cannot access what it is like to think that *P* for a subject *S*, we certainly could access the fact that *S* is

thinking that *P* through neurotechnological mindreading, and this is enough for justifying the legitimacy of the worry about how to protect mental privacy.

When considering mental states as assemblages of properties, there are a number of parts of our mental states to which we seem to have third-person access and which we often would like to keep private. Take, for example, the expression of the contents of our emotions, beliefs, intentions, etc., (what we feel sad about, what we think about a colleague). These are mental events consisting of a subjective character, but also of properties that can be accessed by others and that consequently we often try to conceal. Even if it is true that part of our conscious mental states are only accessed from the first-person perspective, they also have properties accessible from the 3rd person perspective, which can be inferred from neural data. Due to the fundamental role that neural data plays in the instantiation of conscious states in terms of *form* and *content*, we should not underestimate the ways in which access to this type of data could undermine the privacy we have over our own conscious mental states. Although I might feel I have private access to what I'm currently conscious of, this does not mean that others couldn't infer what I'm being aware of through neurotechnological mind-reading. This reinforces the legitimacy of the mental privacy worry.

What about the second argument? Natural mind-reading does not offer direct access to a person's mind or to what it is like to feel something by the person I'm interacting with. However, it does seem to provide indirect or inferential access to the content of her mental states. Mentalizing is exactly that, an ability to infer a person's mental contents (*see*, [62]). The relevant issue here has to do with the kinds of inferences that can be made about others. The ability that other people have to infer our mental states can be often limited by modulating our behavior. We can often change the way we behave or the things we say or express in order to conceal our mental processes. However, technological mind/reading is becoming increasingly powerful to the point of completely undermining this capacity to become *opaque* to others. Even without addressing neurotechnological mind/reading, digital psychological profiling techniques based on behavioral data can be used to infer many different kinds of information about us in a way that we cannot anticipate. Digital profiling can be defined as the use of algorithms to discover

patterns in databases that can be used to represent a group or category and/or the application of this representation to an individual to characterize him or her as a member of a group or category [63]. Profiling techniques that can be applied to digital behavioral data (Facebook likes, tweets, Instagram posts, etc.) provide access to increasingly private information. Kosinski et al. [64] have shown that digital records of Facebook Likes can be used to automatically and accurately predict a number of highly sensitive personal attributes such as sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, substance abuse, separation from parents, age and gender. Neurotechnological mind-reading could push profiling techniques even further, given that they would provide access to additional variables defining neural processes that may not even have a behavioral correlate.

Secondly, the folk-psychology categories on which our natural mind-reading capacity relies have been contested by cognitive neuroscience. The challenges associated with mapping brain structures onto these psychological categories gave rise to a debate about our cognitive ontologies, that is, about what are the basic categories through which we should conceptualize the mind (see [65]). Some neuroscientists and philosophers argue that the systematic mapping between neural structure and psychological function requires to reconceptualize the latter in a bottom up manner, based on what we know about neurocognitive mechanisms (e.g., [65–67]). If our minds are actually characterized by this revised ontology, then, natural mind-reading may not give us access to the real states, processes and capacities that constitute other minds. That is, the folk-psychological categories (such as the concepts of belief, desire or emotion) that our mind-reading capacity uses to understand other people's minds may be useful for efficient social interaction but may not refer to anything real. These psychological categories may be similar to the concept of color as a mind-independent property of the world, which we may not be able to abandon but nonetheless does not seem to correspond to anything real in the world. By contrast, neurotechnological mind-reading based on this theoretical revision of mental states and processes would be the sole means to breach a mind's privacy, it would be the only kind of mind-reading that provides access to the properties that actually constitute our minds.

Finally, the fact that something is widely widespread in our society or evolutionary history does not make it ethically irrelevant. Humans murder each other since the very beginnings of our species, however, this does not make murder ethically irrelevant. Moreover, the comparison between psychological and neurotechnological mind-reading does not respect the different nature of these two processes. While the former is a biological necessary condition for human interaction, the latter is an artificial form of accessing information about others for different purposes that (with the exception of BCIs that restore communication in impaired patients) are often different from daily interactions. While psychological mind-reading cannot really be controlled by us, neurotechnological mind-reading can. Therefore, different ethical conditions might apply to these two types of mind-reading when considering the degree and type of control we can exert over them.

The Empirical Discussion

Two main criticisms support the idea that we should not even worry about legislating about mental privacy yet on the basis of empirical and technological issues. A first argument focuses on the very possibility of neurotechnological mind reading and proposes that concerns about potential intromissions in mental privacy do not seem to have real grounds (e.g., [26, 68]). It is suggested that decoding the contents of our mental states though neurotechnological devices – namely, determining exactly whether a subject is thinking about a muffin, or whether they are imagining a tree – seems very unreal. The idea is that the experimental settings that feed the mental privacy concern consist of studies that allow the observer to distinguish a subject's mental content in a restricted multiple-choice design where both the subject and the decoding algorithm only can choose from predetermined options [26]. Such settings might then be too far from contextualized and ecologically complex real situations where the set of possibilities is not clearly defined. Furthermore, decoding mental states in these settings would depend on previous matching measurements of brain activity evoked in the subject by those categories in the past (e.g., [69, 70]). Therefore, the existence of ecological problems and methodological circularity would make it really hard to

produce real-life mind reading in situations lacking such predetermined conditions.

A second criticism addresses the priority of the concern about mental privacy. The idea is that profiling techniques applied to non-neural data pose a greater risk to mental privacy in current society. The analysis of databases in social and private networks (banks, medical institutions, and the like) would provide access to very sensitive information about our mental processes, states, and dispositions, and these issues should be the priority in the discussion about privacy. As mentioned above, profiling refers both to the process of 'discovering', through mining techniques, correlations between data in databases that can be used to represent an individual or group and also to the application of these sets of correlated data to represent a subject as a member of a group or category [71]. Some of these techniques, such as "psychological targeting" [72], are aimed at building and applying psychological profiles of data subjects that represent their mental states, processes, character traits, psychiatric conditions, etc. Crucially, not only neural data but also some kinds of behavioral digital data can be used to build psychological profiles. Once these profiles are built, they can be applied to digital data generated by a particular subject (for instance, in social media) for representing them as belonging to some psychological category or group. In these cases, mental information that the subject may want to keep private is obtained from data that may be non-sensitive (e.g. Facebook likes and posts) or that they willingly share, and therefore profiling techniques undermine the informational self-determination or control that constitutes mental privacy. Thus, even if we agree that the privacy of mental information is at risk, it is proposed that the safeguards we need may better be condensed into digital rights instead of specifically regulating neurotechnology through particular prohibitions. Recently, Bublitz [31] has suggested that the protection of mental privacy might be already included in the rights protected by the EU General Data Protection Regulation—GDPR henceforth. The GDPR's special category of sensitive data includes genetic and health data (Article 9 (1) GDPR), and, for the author, much data about the brain (neurodata) or stemming from medical examinations of it (neuroimaging) would be covered by this category. Consequently, with some exceptions, processing of such data is prohibited. Indeed, neurodata appears to be

very close to health data. Health data provide information about the physical or mental health of a person. In turn, by providing information about neural activities that take place in the human body, neurodata can also indirectly convey health-related information. However, they can also convey information about mental processes that take place consciously or unconsciously and which may be closely connected to the identity and personality of the person and give them a special quality compared to traditional health data.

The two aforementioned criticisms make important points about the way in which the debate about how to protect mental privacy has been framed in current literature. However, they do not seem to undermine the need to discuss mental privacy in the context of neurotechnology use. Regarding the second objection, it is true that we should prioritize pressing matters at all levels. However, this does not imply that we should not discuss other problems in parallel. In the same way in which legislation about equal access to health services should be a priority, this by no means implies that we should not discuss regulations about constructing buildings on potentially dangerous sites. Certainly, both issues can be discussed in parallel being equally motivated by the possibility of people getting hurt due to lack of regulation. Also, if mental privacy is understood as a concern that is not limited to control over our neural data, but rather as an interest that could also be threatened by digital technologies that allow us to extract information about mental processes or states by analyzing non-neural data, then psychological targeting and similar techniques provide additional reasons to articulate a right to mental privacy. The key question would then be whether privacy risks associated with neurotechnologies are of the same kind as those associated with digital technologies. As we will argue in Section "What is Unique About Neurotechnological Mind-Reading?", there are good reasons to believe that there are significant differences.

Regarding the first objection, it is highly plausible to think that experimental (ecological and methodological) constraints might be overcome in the future. For example, the development of commercial brain-machine interfaces such Facebook's brain-to-text and Kernel Flow are clear attempts to overcome this. Recently, Elon Musk—Neuralink's CEO—implanted a totally functional BCI chip in a pig, what shows

that it is possible to implant this type of invasive neurotechnological device without harming brain structures, and without altering the animal's normal behavior [36]. During the showcase led by Musk, no behavioral and functional differences were observed between the implanted animal and the animals that had their chip previously removed, showing the reversibility of the product. Crucially, after receiving approval from the US Food and Drug Administration and a hospital ethics board, Neuralink started recruiting patients with quadriplegia, or paralysis in all four limbs, due to cervical spinal cord injury or amyotrophic lateral sclerosis (ALS) for a clinical trial. Although information about the study is lacking, Musk reported during February 2024 that the first person with an implanted chip has recovered from the intervention and can now control a computer mouse "just by thinking".[1]

In the same vein, approaches that overcome experimental setting limitations related to the range of mental contents that can be neurotechnologically decodified seem to be on a good track. As already mentioned, decoding methods being developed by Jack Gallant and his colleagues identify a representational system in the cerebral cortex in which, through a flexible and systematic combination of basic neural activity patterns, thousands of different categories can be codified. By understanding this rich representational capacity, neurotechnological mind-reading may no longer be limited to a small set of stimuli with which a decoding algorithm has been trained. Decoding techniques based on models of this cortical representational system can be used to identify any given stimulus taken from a large set of completely new natural images (e.g., [57–61]).

Even if it is true that real neurotechnological mind reading has not been fully developed, people are trying, and, eventually, they might get it right. The tendency to postpone decision-making until we obtain crucial information that we would like to have in order to make a good decision is understandable. However, at the same time, it would be a case of the delay fallacy in the context of mental privacy [73, 74]. This would go as follows: since we do not know what features neurotechnological mind reading

will have, we should wait for such technologies to be fully developed before deciding how to regulate it. The problem with this way of thinking is that it might be too difficult to modify the behavior of providers and users once commercial neuroapplications are already developed. More importantly, damage to users might have already been done. As suggested by Collingridge [75], the real impact of a technology cannot be easily predicted until the technology is extensively developed and widely used, but at the same time, controlling or changing this technology is difficult once the technology has become socially entrenched. A similar situation seems to be shown by the Facebook Case [40]. After a long period of avoiding discussions about the risks of these types of social networks, governments and users only started to seriously worry about privacy after faults had already been committed. We believe that the responsibility of governments and legal organizations should not only be the creation of a reactive agenda, but also a preventive one, so that when it comes to neurotechnological mind reading it may well be better to make an early decision with incomplete information than to make a more informed decision at a later time – or even when it is too late.

What is Unique About Neurotechnological Mind-Reading?

The analysis offered in the last section shows that skepticism against mental privacy threats is not well justified. However, as mentioned in the introduction, once we agree that neurotechnological developments pose significant threats to mental privacy, a second debate opens up. For conservatives, current laws would already cover some of these threats, or minor specifications should be added to do so (e.g., [30–35]). For liberals, the type of threats to mental privacy posed by commercial neurotechnological applications might not be covered by current legal frameworks or international treaties, and for this reason, new regulatory frameworks should be developed (e.g., [17, 20, 24, 28, 33, 36–38]). The debate between conservatives and liberals is still open in the literature. Here we propose that a way of disentangling the discussion implies stepping back from these two specific approaches and focusing on a preliminary task, namely, an analysis of the risks that

---

[1] https://www.wired.com/story/neuralink-brain-implant-elon-musk-transparency-first-patient-test-trial/

are unique to neurotechnological developments. We believe that we can disambiguate this issue by critically examining the uniqueness of the type of data on which neurotechnological mind-reading is based.

Uniqueness of Neural Data

Insofar as the notion of mental privacy refers to the control a subject has over her neural data, determining what is unique about mental privacy requires understanding what is unique about neurotechnological mind-reading as opposed to other forms of mind-reading. We call neurotechnological mind-reading the decoding of information about a subject's mental states, processes, traits, capacities, etc. through the analysis of data about neural structure, activity and/or function. Thus, unique or special features associated with neural data could account for the special nature of neurotechnological mind-reading. At least 4 special features can be associated with neural data.

First, from a biological point of view, neural data are causally closer to behavior than other kinds of sensitive data, such as genetic data. In this sense, neural data would be much more predictive of behavior than genetic data in experimental and non-experimental contexts. Predicting psychological traits from genetic data analysis is very challenging because these traits are determined by a combination of genes in interaction with personal experience, learning, and a number of features of the subject's environment. By contrast, brain activity, structure and function is the result of genome and experience, being causally closer to psychological traits and behavior, and therefore, offering a more accurate capacity for prediction (Farah et al. 2008). Here it is important to note that a number of genetic predispositions might never manifest themselves in behavior, so the predictive power that genetic data could offer might be very variable [76, 77]. In contrast, any observed behavior depends in part on brain activity, structure, and function, and therefore, neural data can be more easily linked to observed behavior than genetic data.

Secondly, neural data exhibits a higher temporal resolution than other types of sensitive data and, therefore, it provides a much powerful tool for manipulating behavior. Neural data can provide access not only to relatively stable psychological traits and dispositions but also to individual, dynamic and temporally indexed mental states, events and processes.

This makes neural data a unique tool for real-time interaction, including real-time brain-computer interfacing [37].

Thirdly, neural data may be seen as providing a more "direct" access to our mental processes and states than behavioral data, to which many psychological analytic tools are applied. When we obtain mental information from neural data, it seems that we are gathering this information "directly from its source", so to speak. This is because, at least in some experimental settings, one may be able to access mental information without the mediation of behavior [27, 56]. In contrast, behavioral data could imply an additional level of inference when establishing causal conclusions about the subject. This issue about the accessibility of neural data reinforces its uniqueness due to the ways it could allow mind-reading and control of a subject's mental states. Importantly, the very possibility of recording neural data underlying specific mental states production might offer scientists and governments the possibility of not only reading, but also controlling the production of mental states in the minds of regular citizens, a process that has been called "brain-hacking" [56].

Even if such a term sounds dystopic, the recollection of neural data potentially enables a more fine-grained manipulation of behavior than behavioral data. Behavioral data provides information about the environmental features (e.g. stimuli, response modality, etc.) that constitute a cognitive task and that can be manipulated in order to trigger or modulate mental processes and behavior. Neural data provides information about the neural mechanisms underlying behavior. The neural components, activities and organizational properties that constitute neural mechanisms define a much wider set of variables that can be targeted by intervention techniques to modulate our cognitive capacities in more complex and subtler ways [78].

Finally, some kinds of neural data can be regarded not merely as information *about* the brain, but rather, information *in* the brain. The function of activity in some neural structures is to codify information about the external or internal environment of the organism, thus forming neural representations. This means that our brain mechanisms themselves are partially constituted by pieces of information. As a consequence, when we talk about protecting neural information we are in part

talking about protecting the brain itself [19]. This opens up to an interesting epistemological issue, namely, the debate about the distinction between the data's owner (the subject) and the data that is owned (neural data), distinction that this fundamental when legislating about privacy and property in law. If in this case there is a substantial difference regarding the data's relation to the data subject, this could can affect the data subject's rights and interests.

In most kinds of personal data (of the type shared with banks, social networks, and the like) there seems to be a clear boundary between the subject and the information that belongs to her, even if this data represents her sensitive information. However, it does not seem to be exactly the same case with neural data (and genetic data), where the data is part and parcel of how and what the data's owner is at the most fundamental biological and existential level. Neural data is part and parcel of what the person is. The subject is, so to speak, partly constituted by that data. Without assuming a hard physicalist stance of the mind, and without examining the many debates that could arise from this issue, it is fair to say that the boundary between the data's owner and data in the case of neural data seems to be less clear than in the other aforementioned cases. This issue seems to add further importance to the uniqueness of neural data, and certainly, this could impact the way in which mental privacy is thought to work and planned to be protected in existing (conservative view) or novel (liberal view) legal frameworks.

Understanding what is unique about neural data may be relevant for both liberal and conservative views. As we will see, some liberals might want to create a legal framework for the protection of mental privacy by using as a template laws originally thought for the protection of other types of data. Problematically, such data might not have the same features as neural data, and therefore a number of potential risks produced by those special features might not be covered. On the other hand, conservatives might also want to try to extend the scope of laws originally thought for the protection of other types of data, also neglecting the unique nature of neural data and the risks that those special features create.

The Unique Risks of Neurotechnological Mind-Reading

In light of the unique features of neural data, what are the risks posited by neurotechnological development? Conservatives and Liberals agree that current neurotechnological devices could lead up to different threats to mental privacy and the unique features of mental data reinforces the need for a discussion about how to protect it. There are at least five main risks that can be associated with neurotechnological mind-reading.

Firstly, there is the risk of arbitrary discrimination. Let's take the case of employment. There are several mental processes that seem necessary to adequately perform a given job. Furthermore, in some occupations failure in performance could put peoples' lives at risk. For instance, we need to assess whether a truck driver or an airplane pilot can sustain attention for a given period of time. However, there are many capacities that should not be taken into account in providing access to a job position. For instance, being a good teacher may be consistent with attention deficits or some forms of neurodivergence and even emotional dysregulation. If neurocognitive profiles based on neural data begin to be used as a means to make decisions about hiring, promoting or firing employees, then there is the risk that these decisions are made on the basis of irrelevant information, leading to arbitrary discrimination. Why is this risk different from the risks associated with the possibility of using, for instance, genetic or behavioral data for making decisions on hiring, promoting or firing employees? Firstly, given that, as we mentioned above, neural data is much more reliable than genetic data for making inferences about psychological traits, it seems more plausible that the former will be actually and more widely employed for making such decisions, even if the traits are irrelevant to the job. Furthermore, the dynamic nature of neural data means that, unlike genetic data, it can be used to evaluate performance of people at the workplace. By contrast, behavioral data can also be used to reliably infer psychological states or traits. However, given that neural data is (at least potentially) gathered directly from people's brains, without the mediation of behavior, the worry here is that workers would have less control over the data gathering process. For instance, we saw that people can conceal their negative intentions and feelings

towards other co-workers by modulating behavior. If performance in the workplace is evaluated by using neural data, then workers could be penalized and disciplined for antisocial attitudes that have not been exteriorized.

A second and closely related risk is that of neurocognitive stigmatization. If these profiles end up providing differential access to goods and services to different people, then different social status may be associated with them, introducing new forms of stigmatization, similar to those that people with psychiatric conditions suffer nowadays. This risk may be different from the forms of stigmatization that could arise from genetic and behavioral data because the fine-grained nature of neural data could be the basis of new forms of classification and assessment of human performance that go beyond current psychological normative categories. Although current neurodiversity movements aim to better understand the nature and neural basis of neurodivergence for promoting inclusion and preventing stigma, the same knowledge could be covertly employed in a discriminatory manner.

A third concern is that neurotechnological mind-reading could provide an opportunity for forced self-incrimination. Basically, the idea is that providing testimony consists in the voluntary exteriorization of our thoughts. Therefore, if neurotechnology can be used to decode our thoughts and make them publicly available without our consent then, when this information is gathered in the context of a criminal procedure and is incriminatory, then this process could be considered as equivalent to being forced to testify against ourselves. It is considered that this is a threat uniquely posed by neural data because, as we mentioned above, this is the kind of data that can potentially provide access to dynamic mental processes without the mediation of behavior.

The same idea of neural data as that increases the degree of transparency of our mental processes introduces two final risks. Fourthly, we can articulate a concern about forced self-knowledge. The idea in this case is that our mind is constituted by an enormous amount of information (our knowledge, memories, etc.) to which we have a very restricted access. We only have access to the small pieces of information we are conscious about. This restricted access could be sometimes a necessary condition for our mental health and psychological well-being. Think about how careful we have to be in providing access to a person to a traumatic memory of an event in her childhood. If neurotechnological mind reading becomes ubiquitous, our minds could become increasingly transparent to ourselves and we may not have control over which aspects of our minds we have access to, therefore risking our mental well-being. Finally, the potential decreased opacity of our minds brought up by ubiquitous neurotechnology also poses a risk for the freedom of thought or cognitive liberty. It is supposed that our minds are the safe-space in which we have a higher degree of freedom. Most of the things we are not allowed to do are things that we can legitimately fantasize about. This freedom to explore our deepest desires and motivations is for some authors the necessary basis for the development of our identities. However, if our thoughts became increasingly accessible to other people, this could undermine our cognitive liberty, and consequently, our identities [33, 79–81].

Concluding Remarks: Should Unique Threats Induce The Creation of Neurorights?

In this paper we have argued that there are good reasons to think that current limitations of neurotechnological mind-reading should not deter us from addressing its current and potential unique threats to mental privacy. Skeptical approaches to the debate seem to ignore the uniqueness of neural data and the specific types of risks that arise from neurotechnological applications. It is important to see what we have done here as a preliminary step to face the conservative/liberal debate about how to specifically protect mental privacy. This, because the offered risk assessment does not seem to support any of these views directly. It could be the case that such risks can be addressed by current regulations (or a slightly updated version of them) or by new neuro-centered rights depending on how each view is defended. Here, we hope to have provided a first input to disambiguate this more specific debate. Certainly, this debate is still unresolved in the literature. Both conservatives and liberals seem to face different argumentative and practical challenges. In this final section we briefly analyze some problems that arise when novel regulations are created to protect mental privacy and how

this would leave room for more alternatives within the liberal side.

A liberal proposal has been pioneered by the NET in collaboration with the Chilean Government. Considering the general concern about the potential misuses of neurotechnologies in society, the Chilean Senate passed the first ever bill for the creation of Neurorights in October 2020. In its Article 4 ([82], Bulletin N° 13.828–19), the bill states that:

> The use of any system or device, be it neurotechnology, BCI, or other, the purpose of which is to access neuronal activity, invasively or noninvasively, with the potential to damage the psychological and psychic continuity of the person, that is, their individual identity, or with the potential to diminish or damage the autonomy of their free will or decision-making capacity, is prohibited.

This bill is thought to set an important example of jurisprudence for any future discussion of different governments around the world aiming at establishing specific policies and laws to regulate the use of intrusive and non-intrusive neurotechnologies in medical (experimental and non-experimental) and non-medical (military and commercial) contexts. Now, regarding our specific discussion, the NET has recently proposed that mental privacy should be protected by the laws that regulate organ transplantation and donation processes [28]. We call this the *Organic Approach* to mental privacy.

The organic approach to mental privacy consists of two main claims [28]. Firstly, people not only have a right to not be compelled to give up brain data but, crucially, brain data collection requires explicit 'opt-in' authorization. In this sense, '*brain data should not be collected passively or rely on individuals to "opt-out" if they do not wish their data to be collected*' [28], p. 14). Instead, the approach requires data collectors to obtain consent not only for such collection, but also for the ways in which neural data will be used in terms of purpose and time. Secondly, people have a right for their neural data not to be commercially transferred and used. Moreover, collecting and using neural data for commercial purposes is prohibited *regardless of consent status*.

How was this idea articulated within the Chilean approach? The proposal of the 'Future Challenges, Science, Technology and Innovation' Commission of the Chilean Senate consists of a constitutional reform bill ([83], Bulletin 13.827–19) and a bill on Neuroprotection ([82], Bulletin 13.828–19). The first version of the latter includes a specific interpretation of mental privacy. It proposed to understand neural data as organic tissue, eliciting protection from the laws for organ transplantation and donation. Article 7 states that:

> [T]he collection, storage, treatment, and dissemination of neuronal data and the neuronal activity of individuals will comply with the provisions contained in Law No. 19.451 regarding transplantation and organ donation, as applicable, and the provisions of the respective health code.

In line with the organic approach, the special nature of neural data seems then to depend on its connection to physical and mental integrity. Article 1.a of this version of the Neuroprotection bill explicitly affirms that the law aims to '[p]rotect the physical and mental integrity of individuals, *through the protection of the privacy of neuronal data*' (emphasis added). It is worth mentioning that, after thorough discussion, the last version of this bill (still being considered in the Chilean Chamber of Deputies) dropped the organic proposal. In its new Article 11, the bill only states that.

> [N]eural data are, as a general rule, reserved and their collection, storage, processing, communication and transfer shall only be used for the legitimate and informed purposes to which the person has consented, under the terms provided for in this law [the bill establishes prior, explicit and specific informed consent for any neurotechnological application] [...] The neuronal data shall be treated as sensitive data under the terms of Law No. 19,628, on the protection of privacy.

However, the organic approach was later articulated by Brazilian Bill nº 1229/21 [84] aimed at protecting neural data. While the Chilean proposal consists of an autonomous bill for dealing with neuroprotection in general, the Brazilian bill (pending in the Chamber of Deputies), was an amendment of the Brazilian General Law on Personal Data Protection (LGPD). The Brazilian proposal defined neural data as a special kind of information in need of special

protection, stronger than that provided for other kinds of sensitive data (Article 13-E and "Justification"). As to restricting the commercialization of neural data, Brazil seemed to follow the original proposal of Chile. Article 13-C affirmed that 'communication or shared use between controllers of neural data with the objective of obtaining economic advantage is forbidden'. Although the connection between mental privacy and other rights within this regulation has not been made explicit, the Brazilian proposal does recognize neural data subjects to be vulnerable in a way that goes beyond usual privacy concerns. Article 13-B states that 'it is forbidden to use any brain-computer interface or method that may cause damage to the data subject's individual identity, impair his or her autonomy or psychological continuity'. Although this approach might be regarded as *prima facie* practical and intuitive, we shall point out some difficulties that should be considered before trying to include it into a specific legal framework.

The analogy seems to articulate the intuition mentioned above that neural data is not merely information about our brains but rather a constitutive part of them. However, we think that this way of articulating this intuition has several problems. As Wajnerman-Paz [19] suggested, a key difference between neural data extraction and bodily organ harvesting is that the former does not involve any organic material transfer. A biopsy is a typical example of information transfer involving the manipulation of organic material. In this case, data is contained in organic tissue that needs to be extracted in order to be analyzed. By contrast, in neural data extraction the physical integrity of the subject is not necessarily affected because information can be gathered without extracting any organic material. In this case, information is strictly speaking not transferred but rather 'replicated' [85]. That is, we do not move an information-carrying object but rather produce a new object (a copy) carrying the same information at a different point in a communication network. For instance, EEG signals that carry neural data are not constituted by neural tissue but rather by patterns of electrons in EEG electrodes. They are a copy of physically different signals contained in neural waves of ions, with which EEG electrodes interact. How is it then conceptually precise to use an organic-based approach to protect an entity that does not have any organic element? In making such a comparison, we might lose some of the most

relevant aspects of the type of entity we are trying to protect. The focus of protection in all circumstances is on the data subject, which means that the way in which some kind of data is or is not (ontologically or epistemologically) close to the identity of the data subject is critical to assessing whether it constitutes sensitive information. Unless we provide an appropriate characterization of the relation between the subject and its data, we cannot determine whether neural data can be placed under the existing data protection categories, constructed on the basis of the risk posed by their processing to the rights and interests of the data subject. Thus, it could be that the more adequate way to protect neural data is by regarding it a special kind of sensitive data (such as genetic data) whose protection requires reinforced measures specified in general data protection regulations.

It may be true that neural data is ontologically closer to the ultimate nature of ourselves than other kinds of data. However, this feature cannot ground the organic analogy. More generally, the idea that neural data is unique and entails unique threats, is only a first step that legitimates or substantiates the discussion between conservatives and liberals. Only if there are such unique risks, it makes sense to examine whether current regulations can address them appropriately. However, uniqueness claims do not tilt the balance toward liberal views. It is possible that current regulations are enough to address risks that were not foreseen when they were drafted.

Declarations

Conflict of Interest   The authors have no relevant financial or non-financial interests to disclose.

References

1. Andrews, A., and P. Weiss. 2012. Nano in the brain: Nano-neuroscience. *ACS Nano* 23 (10): 8463–8464.
2. Alivisatos, A.P., M. Chun, G.M. Church, K. Deisseroth, R. Greenspan, P.R.J. McEuen, M.L. Roukes, T.S. Sejnowski, P. Weiss, and R. Yuste. 2013. The brain activity map. *Science* 339: 1284–1285.

3. Alivisatos, A.P., M. Chun, G.M. Church, et al. 2015. A national network of neurotechnology centers for the brain initiative. *Neuron* 88 (3): 445–448.

4. Koch, C., and R.C. Reid. 2012. Neuroscience: Observatories of the mind. *Nature* 483: 397–398.

5. Quaglio, G., P. Toia, and E. Moser. 2021. The International Brain Initiative: Enabling Collaborative Science. *The Lancet Neurology* 20 (12): 985–986.

6. Chase, H.W., M.A. Boudewyn, C.S. Carter, et al. 2020. Transcranial direct current stimulation: A roadmap for research, from mechanism of action to clinical implementation. *Molecular Psychiatry* 25 (2): 397–407.

7. Chaudhary, U., N. Birbaumer, and A. Ramos-Murguialday. 2016. Brain-computer interfaces for communication and rehabilitation. *Nature Rev Neurol.* 12 (9): 513–525.

8. Espay, A.J., P. Bonato, F. Nahab, et al. 2016. Technology in Parkinson disease: Challenges and opportunities. *Movement Disorders* 31 (9): 1272–1282.

9. Garnaat, S.L., S. Yuan, H. Wang, et al. 2018. Updates on transcranial magnetic stimulation therapy for major depressive disorder. *Psychiatric Clinics of North America* 41 (3): 419–431.

10. McFarland, D.J. 2020. Brain-computer interfaces for amyotrophic lateral sclerosis. *Muscle and Nerve* 61 (6): 702–707.

11. Ababkova, M. Yu., Pokrovskaia, N. N., and Tros-tinskaya, I. R. 2018. Neuro-Technolo-gies for Knowledge Transfer and Expe-rience Communication. *The European Proceedings of Social & Behavioural Sciences (EpSBS)*. XXXV, 10–18.

12. Doya, K., A. Ema, H. Kitano, M. Sakagami, and S. Russell. 2022. Social impact and governance of AI and neurotechnologies. *Neural networks : The official J Int Neural Netw Soc.* 152: 542–554.

13. Fernandez, A., N. Sriraman, B. Gurevitz, et al. 2015. *Pervasive neurotechnology: A ground-breaking analysis of 10,000+ patent filings transforming medicine, health, entertainment and business*. San Francisco: SharpBrains.

14. López-Silva, P., and L. Valera. 2022. *Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights*. Amsterdam: Springer.

15. López-Silva, P., and R. Madrid. 2021. Sobre la conveniencia de incorporar los neuroderechos en la constitución o en la ley. *Revista Chil De Derecho y Tecnol* 10 (1): 53–76.

16. Privitera A.J, and Du, H. 2022. Educational neurotechnology: Where do we go from here?, Trends in Neuroscience and Education, Volume 29, 100195, https://doi.org/10.1016/j.tine.2022.100195.

17. Ienca, M., and R. Andorno. 2017. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sci Soc Policy* 13 (1): 1–27.

18. Lavazza, A. 2018. Freedom of thought and mental integrity: The moral requirements for any neural prosthesis. *Frontiers in Neuroscience* 12: 82.

19. Wajnerman-Paz, A. 2021a. Is your neural data part of your mind? Exploring the conceptual basis of mental privacy. *Minds Mach.* https://doi.org/10.1007/s11023-021-09574-7.

20. Ienca, M. 2021. On Neurorights. *Front Human Neurosci.* 15. https://doi.org/10.3389/fnhum.2021.701258.

21. Shen, F.X. 2013. Neuroscience, mental privacy, and the law. *Harvard J Law Publ Policy* 36: 653.

22. Wajnerman-Paz, A. 2021b. Is Mental Privacy a Component of Personal Identity?. *Frontiers in Human Neuroscience* 604. https://doi.org/10.3389/fnhum.2021.773441.

23. Wajnerman-Paz, A., and P. López-Silva. 2022. Mental Privacy and Neuroprotection: An Open Debate. In *Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights*, ed. P. López-Silva and L. Valera, 144–155. Amsterdam: Springer.

24. Yuste, R., J. Genser, and S. Herrmann. 2021. It's Time for Neuro-Rights. *Horizons: J Int Relations Sustain Dev.* 18: 154–16.

25. Gilead, A. 2015. Can brain imaging breach our mental privacy? *Rev Philo Psychol.* 6 (2): 275–291.

26. Meynen, G. 2019. Ethical issues to consider before introducing neurotechnological thought apprehension in psychiatry. *AJOB Neuroscience* 10 (1): 5–14.

27. Ryberg, J. 2017. Neuroscience, mind reading and mental privacy. *Res Publica* 23 (2): 197–211.

28. Göering, S., E. Klein, L. Sullivan, et al. 2021. Recommendations for responsible development and application of neurotechnologies. *Neuroethics* 14: 365–386. https://doi.org/10.1007/s12152-021-09468-6.

29. Göering, S., and R. Yuste. 2016. On the necessity of ethical guidelines for novel neurotechnologies. *Cell* 167 (4): 882–885.

30. Borbón, D., and L. Borbón. 2021. A Critical Perspective on NeuroRights: Comments Regarding Ethics and Law. *Front Human Neurosci* 15: 703121. https://doi.org/10.3389/fnhum.2021.703121.

31. Bublitz, J.C. 2022. Novel Neurorights: From Nonsense to Substance. *Neuroethics* 15: 7. https://doi.org/10.1007/s12152-022-09481-3.

32. Hertz, N. 2023. Neurorights – Do we Need New Human Rights? A Reconsideration of the Right to Freedom of Thought. *Neuroethics* 16: 5. https://doi.org/10.1007/s12152-022-09511-0.

33. Ligthart, S., M. Ienca, G. Meynen, F. Molnar-Gabor, R. Andorno, C. Bublitz, P. Catley, L. Claydon, T. Douglas, N. Farahany, J. Fins, S. Goering, P. Haselager, F. Jotterand, A. Lavazza, A. McCay, A. Wajnerman Paz, S. Rainey, J. Ryberg, and P. Kellmeyer. 2023. Minding rights: Mpping ethical and legal foundations of 'neurorights.' *Cambridge Quarterly of Healthcare Ethics* 2023: 1–21.

34. Susser, D., and L.Y. Cabrera. 2023. Brain data in context: Are new rights the way to mental and brain privacy?. *AJOB neuroscience,* 1–12.

35. Zuñiga-Fajuri, A., L. Villavicencio, and D. Zaror-Millares. 2021. Neurorights in Chile: Between neuroscience and legal science. In *Regulating Neuroscience: Transnational Legal Challenges*, ed. M. Hevia, 165–179. Amsterdam: Elsevier.

36. Baselga-Garriga, C., P. Rodriguez, and R. Yuste. 2022. *Neuro rights: A human rights solution to ethical issues of neurotechnologies*, 157–161. Protecting the Mind: Springer.

37. Ienca, M., Fins, J. J., Jox, R. J., Jotterand, F., Voeneky, S., Andorno, R., ... & Kellmeyer, P. 2022. Towards a governance framework for brain data. *Neuroethics*, 15(2), 20. https://doi.org/10.1007/s12152-022-09498-8.

38. Yuste, R., S. Göering, and B.A.Y. Arcas. 2017. Four Ethical Priorities for Neurotechnologies and AI. *Nature* 551 (7679): 159–163.

39. Presidential Commission for the Study of Bioethical Issues, 2014. Gray matters. In *Integrative approaches for neuroscience, ethics and society, vol 1*. Bioethics Commission, Washington, DC.

40. López-Silva, P. 2022. The Concept of Mind in the Neuroprotection Debate. In *Protecting the Mind: Challenges in Law, Neuroprotection, and Neurorights*, ed. P. López-Silva and L. Valera, 9–18. Amsterdam: Springer.

41. Bublitz, J.C. 2013. My mind is mine!? Cognitive liberty as a legal concept. In *Cognitive Enhancement,* ed. Franke, H. 233–264. Amsterdam: Springer. https://doi.org/10.1007/978-94-007-6253-4_19.

42. Farahany, N.A. 2012. Incriminating Thoughts. *Stanford Law Review* 64: 352–408.

43. Ligthart, S. 2023. Mental Privacy as Part of the Human Right to Freedom of Thought?. Forthcoming in M. Blitz and JC Bublitz (eds.), The Law and Ethics of Freedom of Thought, 2.

44. Nawrot, O. 2019. What about the interior castle? Response to Ienca's and Andorno's new human rights in the age of neuroscience and neurotechnology. *Roczniki Teologiczne* 66: 69–85. https://doi.org/10.18290/rt.2019.66.3-5.

45. Sententia, W. 2004. Neuroethical considerations: Cognitive liberty and converging technologies. *Annals of the New York Academy of Sciences* 1013: 223.

46. López-Silva, P., and R. Madrid. 2022. Protegiendo la Mente: Un análisis al Concepto de lo Mental en la Ley de Neuroderechos. *Revista de Humanidades de Valparaíso* (20): 101–117.

47. López Silva, P., Madrid, R. 2022b. Acerca de la protección constitucional de los neuroderechos : la innovación chilena [en línea]. Prudentia Iuris. 2022, 94. https://doi.org/10.46553/prudentia.94.2022.pp.39-68. Disponible en: https://repositorio.uca.edu.ar/handle/123456789/15662

48. Muñoz, J.M. 2019. Chile–right to free will needs definition. *Nature* 574 (7780): 634–635.

49. Wellman, C. 1999. *The Proliferation of Rights: Moral Progress or Empty Rhetoric?* London: Routledge.

50. Margulis, S.T. 2003. On the status and contribution of Westin's and Altman's theories of privacy. *J Soc Issues* 59 (2): 411–429.

51. Jackson, F. 1986. What Mary didn't know. *The Journal of Philosophy* 83 (5): 291–295.

52. Kriegel, U. 2015. *The Varieties of Consciousness*. Oxford: OUP.

53. Nagel, T. 1974. What is it like to be a bat? *Philosophical Review* 83: 435–456.

54. Mahy, C.E., L.J. Moses, and J.H. Pfeifer. 2014. How and where: Theory-of-mind in the brain. *Dev Cognitive Neurosci.* 9: 68–81.

55. Miller, S. 2022. *Advanced theory of mind*. Oxford: OUP.

56. Yuste, R. 2020. Si puedes leer y escribir la actividad neuronal, puedes leer y escribir las mentes de la gente. *El País,* December 4th, 2020. https://elpais.com/retina/2020/12/03/tendencias/1607024987_022417.html.

57. Kay, K.N., T. Naseralis, and E. Prengler. 2008. Identifying natural images from human brain activity. *Nature* 452 (7185): 352–355.

58. Huth, A.G., S. Nishimoto, and A. Vu. 2012. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron* 76 (6): 1210–1224.

59. Kay, K.N., T. Naselaris, R.J. Prenger, and J.L. Gallant. 2008. Identifying natural images from human brain activity. *Nature* 452 (7185): 352–355.

60. Huth, A.G., T. Lee, S. Nishimoto, et al. 2016. Decoding the semantic content of natural movies from human brain activity. *Frontiers Syst Neurosci* 10: 81. https://doi.org/10.3389/fnsys.2016.00081.

61. Wen, H., J. Shi, and Y. Zhang. 2018. Neural encoding and decoding with deep learning for dynamic natural vision. *Cerebral Cortex* 28 (12): 4136–4160.

62. Apperly, I. 2010. *Mindreaders*. London: Psychology Press.

63. Custers, B.H.M. 2012. Predicting data that people refuse to disclose; how data mining predictions challenge informational self-determination. *Privacy Observatory Magazine* 3: 1–3.

64. Kosinski, M., D. Stillwell, and T. Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences* 110 (15): 5802–5805.

65. Price, C.J., and K.J. Friston. 2005. Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology* 22 (3): 262–275.

66. Anderson, M. L. (2021). *After phrenology: Neural reuse and the interactive brain*. MIT Press.

67. Poldrack, R.A., and T. Yarkoni. 2016. From brain maps to cognitive ontologies: Informatics and the search for mental structure. *Annual review of psychology* 67: 587–612.

68. Zuk, P., and G. Lázaro-Muñoz. 2019. Ethical analysis of "Mind Reading" or "Neurotechnological Thought Apprehension": Keeping potential limitations in mind. *AJOB neuroscience* 10 (1): 32–34.

69. Cox, D.D., and R.L. Savoy. 2003. Functional magnetic resonance imaging (fMRI) "Brain Reading": Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage* 19 (2): 261–270.

70. Haxby, J.V., M.I. Gobbini, M.L. Furey, et al. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293 (5539): 24252430.

71. Hildebrandt, M. 2008. Defining profiling: a new type of knowledge?. In Profiling the European citizen (pp. 17–45). Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6914-7_2.

72. Matz, S.C., R. Appel, and M. Kosinski. 2020. Privacy in the age of psychological targeting. *Curr Opinion Psychol.* 31: 116–121.

73. Hansson, S.O. 2004. Fallacies of risk. *J Risk Res* 7 (3): 353–360.

74. Mecacci, G., and P. Haselager. 2019. Five criteria for assessing the implications of NTA technology. *AJOB Neuroscience* 10 (1): 21–23.

75. Collingridge, D. 1982. *The* social control of technology. New York: St. Martin's Press.

76. Chapman, C.R., K.S. Mehta, B. Parent, and A.L. Caplan. 2020. Genetic discrimination: Emerging ethical challenges in the context of advancing technology. *Journal of Law and the Biosciences* 7 (1): 1.

77. Prince, A.E., and B.E. Berkman. 2012. When does an illness begin: Genetic discrimination and disease manifestation. *The J Law, Med Ethics* 40 (3): 655–664.

78. Craver, C. F. 2007. *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Clarendon Press.

79. Bublitz, J.C. 2019. Privacy concerns in brain–computer interfaces. *AJOB neuroscience* 10 (1): 30–32.

80. McCarthy-Jones, S. 2019. The autonomous mind: The right to freedom of thought in the twenty-first century. *Front Artif Intell.* 2: 19.

81. Rainey, S., S. Martin, A. Christen, P. Mégevand, and E. Fourneret. 2020. Brain recording, mind-reading, and neurotechnology: Ethical issues from consumer devices to brain-based speech decoding. *Sci Eng Ethics* 26: 2295–2311.

82. Senate of Chile. 2020a. Bulletin N°13.828–19. Available in: https://www.diarioconstitucional.cl/wp-content/uploads/2020/12/boletin-13828-19-nuroderechos.pdf.

83. Senado de Chile. 2020b. Boletín 1382719. Available in: https://www.senado.cl/appsenado/templates/tramitacion/index.php?boletin_ini=13827-19.

84. Câmara dos Deputados do Brasil. 2021. Projeto de Lei 1229/21 (Modifica a Lei n° 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais), a fim de conceituar dado neural e regulamentar a sua proteção.). Available at: https://www.camara.leg.br/propostas-legislativas/2276604.

85. Borgatti, S.P. 2005. Centrality and network ow. *Social Networks* 27 (1): 55–71.