



A Conceptual Framework to Safeguard the Neuroright to Personal Autonomy

José M. Muñoz · Javier Bernácer ·
Francisco Güell

Received: 1 June 2022 / Accepted: 15 June 2023
© The Author(s) 2023

Abstract In this article, we propose a philosophical exploration on the main problems involved in two neurorights that concern autonomous action, namely free will and cognitive liberty, and sketch a possible solution to these problems by resorting to a holistic interpretation of human actions. First, we expose the main conceptual and practical issues arising from the neuroright to “free will,” which are far from minor: the term itself is denied by some trends participating in the neurorights debate, the related concept of ultimate control is also disputed, the understanding of free will depends on cultural context, and the exercise of being free to act in several domains may be covered by other regulations. Second, we analyze the historical origin of cognitive liberty, its current status, and its relation with free will. Third, we criticize the concept of decision in mainstream action theory and propose to conceive action as a unified process constituted by three explanatory, non-sequential dimensions: intention,

decision, and action realization. Fourth, we discuss two possible cases involving neurotechnologies and suggest ways to interpret them according to a unified framework in which free will and cognitive liberty fall under a single neuroright to personal autonomy. Finally, we outline a recommendation to introduce freedom of thought and personal autonomy as complementary neurorights to protect both the internal and external dimensions of thought and action.

Keywords Neurorights · Free will · Cognitive liberty · Decision · Personal autonomy · Freedom of thought

“Freedom, Sancho, is one of the most precious gifts that the heavens gave to men; the treasures that the earth holds and the sea conceals cannot compare to it: for freedom, as well as for honor, one can and must venture one’s life; and, on the contrary, captivity is the greatest evil that can come to men.”

*Don Quixote*¹

¹ Original text (in Spanish): “La libertad, Sancho, es uno de los más preciosos dones que a los hombres dieron los cielos; con ella no pueden igualarse los tesoros que encierra la tierra ni el mar encubre: por la libertad, así como por la honra, se puede y debe aventurar la vida; y, por el contrario, el cautiverio es el mayor mal que puede venir a los hombres” (Miguel de Cervantes, *Don Quixote*, Chapter LVIII). http://www.cervantesvirtual.com/obra-visor/el-ingenioso-hidalgo-don-quijote-de-la-mancha--0/html/fe04e52-82b1-11df-acc7-002185ce6064_20.html#I_153_. Accessed 18 Mar 2023. Translated to English by us.

J. M. Muñoz (✉) · J. Bernácer · F. Güell
Mind-Brain Group, Institute for Culture and Society (ICS),
University of Navarra, Pamplona, Spain
e-mail: jmunozorteg@unav.es

J. M. Muñoz
Gordon Center for Medical Imaging, Massachusetts
General Hospital, Harvard Medical School, Boston, MA,
USA

J. M. Muñoz · J. Bernácer
International Center for Neuroscience and Ethics (CINET),
Tatiana Foundation, Madrid, Spain

Introduction: Neurotechnology, Human Rights, and Autonomy

Even though the president of the United States at the time, George H. W. Bush, proclaimed the 1990s to be the “Decade of the Brain,”² it is instead in the last 10 to 15 years that we have witnessed the most outstanding advances in the history of neurotechnology, that is, “any technologies used to investigate, modulate, repair, or improve the nervous system and its functioning” [1]. The techniques used for this include brain-computer interfaces (see [2]), brain imaging (see [3]), trans- and intracranial stimulation (see [4]), neuropharmaceuticals (see [5]), and brain implants, among others. These advances have opened the door to thinking about promising treatments that may emerge soon for different types of neurological and psychiatric disorders. However, they have also raised concerns regarding eventual misuse or malicious use of neurotechnologies that could undermine certain fundamental rights. For example, experiments with higher mammals have shown that it can be relatively easy to change movement choices in macaques by ultrasound waves [6] and to induce highly aggressive, predatory behaviors in rodents by optogenetic stimulation [7], so such technologies’ application to humans, even though it might seem far-fetched nowadays, could have clear implications for freedom and autonomy.

In light of these and other findings related to agency, identity, privacy, and equal opportunities, several scholars have raised the idea that the current human rights framework needs to be revised, as it is insufficient to address the new challenges posed by neurotechnology. To do this, they maintain, it would be appropriate to introduce new principles or *ad hoc* rights, called *neurorights*, which “can be defined as the ethical, legal, social, or natural principles of freedom or entitlement related to a person’s cerebral and mental domain; that is, the fundamental normative rules for the protection and preservation of the human brain and mind” [8]. The new rights to be included vary according to different proponents of neurorights, who are mainly divided into two groups of researchers. On the one hand, Ienca and Andorno [9] propose rights to *cognitive liberty*, mental privacy, mental

integrity, and psychological continuity. On the other hand, the NeuroRights Foundation [10], Yuste et al. [11] and the Neurotechnology Ethics Taskforce—*a.k.a.* the “Morningside Group”—suggest incorporating *free will*, personal identity, mental privacy, fair access to mental augmentation, and protection from algorithmic biases [12, 13]. Remarkably, most of these neurorights are formulated as rights to keep a human faculty—*i.e.*, a characteristic that humans already have, which is already protected in the Universal Declaration of Human Rights (UDHR)—*despite* neurotechnologies. In other words, they are *negative rights* intended to protect users from unwanted and malicious uses of neurotechnologies.

These proposals have already had political and regulatory influence in several countries, leading to constitutional reform in Chile,³ as well as article 19.2 of France’s new bioethics law,⁴ a brain data protection bill in Brazil,⁵ a procedural law bill in Argentina,⁶ and digital rights charters in Spain⁷ and Mexico.⁸ These proposals are also having consequences at regional and global levels, as demonstrated by recent declarations and resolutions by the Inter-American Juridical Committee,⁹ the Latin American Parliament (*Parlatino*),¹⁰ and (remarkably) the United Nations Human Rights Council.¹¹

Autonomy can be defined as “an individual’s capacity for self-determination or self-governance” and includes moral autonomy (“the capacity to deliberate and to give oneself the moral law”), political autonomy (“the property of having one’s decisions

² <https://www.loc.gov/loc/brain/proclaim.html>

³ <https://www.diariooficial.interior.gob.cl/edicionelectronica/index.php?date=25-10-2021&edition=43086-B&v=2>

⁴ <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT00004384384>

⁵ <https://www.camara.leg.br/propostas-legislativas/2276604>

⁶ <https://www.hcdn.gob.ar/proyectos/proyecto.jsp?exp=0339-D-2022>

⁷ <https://www.lamoncloa.gob.es/presidente/actividades/Paginas/2021/140721-derechos-digitales.aspx>

⁸ https://www.infocdmx.org.mx/doctos/2022/Carta_DDigital.pdf

⁹ https://www.oas.org/es/sla/cji/temario_actual_Desarrollo_estandares_internacionales_neuro_derechos.asp

¹⁰ <https://parlatino.org/comision-de-seguridad-ciudadana-combate-y-prevencion-al-narcotrafico-terrorismo-y-crimen-organizado/declaracion-neuroderechos/>

¹¹ <https://undocs.org/A/HRC/RES/51/3>

respected, honored, and heeded within a political context”), and personal autonomy, that is, “the capacity to decide for oneself and pursue a course of action in one’s life” [14]. In this article, we explore two critical neurorights that concern *personal autonomy*: free will (as proposed by the NeuroRights Foundation [10]) and cognitive liberty (as proposed by Ienca and Andorno [9], among others). Although there are more neurorights related to autonomy proposed by other authors (i.e., mental self-determination [15] and mental freedom [16]), we will concentrate on the two mentioned above because they have had arguably the most normative and media resonance to date.

To do so, and with regulatory applications in mind, we propose a philosophical conceptual exploration of the challenges that proponents of neurorights about personal autonomy should face. Our main goal is to expose the main problems of the concepts involved in these neurorights (free will, cognitive liberty) and autonomous action. Also, we sketch a possible solution to these problems by resourcing to non-controversial terms and a holistic interpretation of human actions. We intend to enrich and focus the debate on these particular neurorights, to find a framework where the many disciplines involved in this topic—neuroscience, law, philosophy, psychology, and engineering, among others—may find common ground wherein a solid proposal may be built.

Sections “[The Longstanding Problem with Free Will](#)”, “[The Neuroright to Cognitive Liberty: How Does it Relate to Free Will?](#)”, and “[Decision, Intention, Action Realization: The Triple Dimension of Action](#)” of this manuscript expose the main conceptual and practical problems arising from the neurorights focused on personal autonomy. In Section “[The Longstanding Problem with Free Will](#)”, we present four challenges related to the concept of free will, which pose severe difficulties in defining it as a neuroright. In Section “[The Neuroright to Cognitive Liberty: How Does it Relate to Free Will?](#)”, we follow a similar approach to the concept of cognitive liberty, which is related to the preservation of one’s own mental states. Section “[Decision, Intention, Action Realization: The Triple Dimension of Action](#)” aims to discuss the difficulty in disentangling the different elements of actions, which could be fundamental to discuss free will, cognitive liberty, and personal autonomy. What action elements (intentions, decisions, action realizations) should be in the spotlight in these debates?

We expose how mainstream action theories do not disambiguate these concepts; as a matter of fact, we recommend understanding them as different dimensions of the same unified process. In Section “[Toward a Neuroright to Personal Autonomy](#)”, considering the challenges in the previous sections, we discuss two possible cases involving neurotechnologies and suggest ways to interpret them according to a unified framework in which free will and cognitive liberty fall under a single *neuroright to personal autonomy*. In the final section, we outline a recommendation to introduce freedom of thought and personal autonomy as complementary neurorights to protect both the internal and external dimensions of thought and action.

The Longstanding Problem with Free Will

The NeuroRights Foundation [10] proposes to incorporate the right to *free will* (FW). It defines it as follows: “Individuals should have ultimate control over their own decision making, without unknown manipulation from external neurotechnologies.” Incorporating FW as a new right—either through a possible reform of the Universal Declaration of Human Rights (UDHR) [17] or through legal systems at the national level—would entail essential difficulties that should not be ignored [18, 20]. It should be considered, as we mentioned in the [Introduction](#), that this is not conceived as a right to achieve FW but as a right to keep on having FW despite neurotechnologies. Freedom of will and choice constitutes an essential ethical principle that underpins many of the world’s legal systems, especially in the West. In keeping with this tradition, it seems appropriate to try to trace a study route for the right to FW. An *ad hoc* UN expert committee has been proposed to work “to develop an international consensus definition of neuro-rights” [11], and a recent proposal to seek a minimal definition of the neuroright to FW aligns with this approach [18].

There is a long story of controversies around FW in neuroscience and philosophy, among other fields of knowledge. Even though we align with those defending the importance of FW in any topic involving human action, this is not the case for many scholars. Being so, it looks futile to find common ground to discuss the neuroright to FW with academics that refuse the existence of FW.

Thus, the first challenge for the neuroright to FW is the *philosophical multidimensionality* of the concept of FW. The relationship between FW and determinism—whether they are compatible or not—is an unresolved problem in the Western philosophical tradition. A diverse spectrum of positions can be found [21]:

- Compatibilism: Determinism and FW are compatible.
- Incompatibilism: Determinism and FW are incompatible. This includes three different sub-positions:
 - Libertarianism: Determinism is not true, so we have FW.
 - Hard determinism: Determinism is true, so we do not have FW.
 - Hard incompatibilism: Both determinism and indeterminism are incompatible with FW, so we do not have FW.
- Disruptive positions that reject the compatibilism-incompatibilism dichotomy, such as revisionism, subjectivism, and agnostic autonomism.

Considering the longstanding discussion about the topic, it would be more practical to overlook the term FW and concentrate on what features of human actions make them autonomous, and thus in which sense a human agent can be considered responsible for their actions.

The second challenge is the inclusion of the concept of *ultimate control* as a central element of the NeuroRights Foundation’s definition of FW: “Individuals should have ultimate control over their own decision making, without unknown manipulation from external neurotechnologies” [10]. An agent’s ultimate control, that is, one’s ability to establish oneself as the genuine author and ultimate source of one’s decisions and/or actions [21], is seen as a necessary condition for FW among libertarians, but this has been criticized by other philosophers (e.g., [22]). One of the main criticisms raised against ultimate control (or “ultimate responsibility”) is that it leads to an infinite regress, as Galen Strawson argues (as described by O’Connor and Franklin [23], §3.1):

“Strawson associates free will with being ‘ultimately morally responsible’ for one’s actions.

He argues that, because how one acts is a result of, or explained by, “how one is, mentally speaking” (*M*), for one to be responsible for that choice one must be responsible for *M*. To be responsible for *M*, one must have chosen to be *M* itself—and that not blindly, but deliberately, in accordance with some reasons r_1 . But for that choice to be a responsible one, one must have chosen to be such as to be moved by r_1 , requiring some further reasons r_2 for such a choice. And so on, *ad infinitum*. Free choice requires an impossible infinite regress of choices to be the way one is in making choices.”

Once again, if including the concept of ultimate control in the neurorights involves endorsing libertarianism, this will end up in controversies sooner or later. Therefore, seeking a definition that reconceptualizes or dispenses with ultimate control seems sensible.

The third challenge for the neuroright to FW relates to its *cultural contextualization*. Several studies show that FW is a culturally variable concept and that Western societies prioritize FW as a value more than Eastern societies. For example, Chernyak and collaborators [24] observe that, while Singaporean children “were more likely to elaborate on lack of free will by referencing punishment and/or having to seek permission from authorities,” U.S. children were inclined “to endorse the freedom to act against norms.” Other works found that various “lexical expressions of ‘free will’ [i.e., in Lithuanian, Hindi, Chinese, and Mongolian] do not refer to the same concept of free will” [25] and also that dualistic intuitions—according to which our “selves” are each the result of a brain and a non-physical mind—incline us to believe in FW [26]. As suggested by Herrera-Ferrá et al. [19]: “These findings invite us to consider culturally shaped factors such as social norms, languages, and beliefs as fundamental factors within an international debate aimed [at finding] a consensus definition of neurorights in general, and free will in particular.” Following this approach, the formulation of neurorights to guarantee personal autonomy must be sensitive to diverse conceptual conceptions of the topic.

The fourth challenge is the *regulatory fitting* of FW. It is essential to try to ensure that introducing FW as a new right is not redundant within the current provisions to avoid overregulation. In the UDHR

[17], for instance, we can find several articles in which the freedom to decide or act is contemplated for different ordinary situations. Here are some examples (the italics are ours):

- Art. 16.2 (free marriage): “Marriage shall be entered into only with the *free and full consent* of the intending spouses.”
- Art. 18 (free thought): “Everyone has the right to *freedom of thought, conscience and religion*; this right includes freedom to change [their] religion or belief, and freedom, either alone or in community with others and in public or private, to manifest [their] religion or belief in teaching, practice, worship and observance.”
- Art. 21.3 (free voting): “The *will of the people* shall be the basis of the authority of government; this will shall be expressed in periodic and genuine elections which shall be by universal and equal suffrage and shall be held by secret vote or by equivalent *free voting* procedures.”

Therefore, it is essential to detect what legal coverage—not only in the UDHR but also in regional treaties and national regulations—already guarantees, in one way or another, freedom of choice and action. In this sense, the line of action would be to determine how implementing FW as a neuroright can effectively expand current protections while avoiding overregulation.

It is interesting to note that the composition of the UDHR articles above seem to fit well with a specific position on the philosophical problem of FW, i.e., so-called classical compatibilism, according to which “*being free* means having the ability to do what we want, desire, or choose to do and [also] that there is no [coercion or constraint] that prevents us from exercising this capacity” [21].¹² However, FW as such is not mentioned.

In conclusion, the difficulties in incorporating a neuroright labeled “free will” are far from minor: the term itself is denied by some trends participating in the neurorights debate, the related concept of ultimate control is also disputed, the understanding of FW

depends on cultural context and an excessive “West-erncentrism” must be avoided, and other regulations may cover the exercise of being free to act in several domains. These challenges should be considered in the general framework where neurorights pertaining to FW are to be discussed. In the following section, we will similarly analyze the neuroright to cognitive liberty and its conceptual similarity concerning FW.

The Neuroright to Cognitive Liberty: How Does it Relate to Free Will?

If, as we suggested in Section “Introduction: Neurotechnology, human rights, and autonomy”, personal autonomy can be understood as “the capacity to decide for oneself and pursue a course of action in one’s life” [14], then FW is not the only proposed neuroright that has been conceived to protect personal autonomy. The other is that of *cognitive liberty* (CL), whose origin dates back to the very beginning of this century, when, in a four-part essay published in the *Journal of Cognitive Liberties*, Boire [27] focused on “each individual’s fundamental right to control his or her own consciousness.” A few years later, Sententia [28] defined CL as

“every person’s fundamental right to think independently, to use the full spectrum of his or her mind, and to have autonomy over his or her own brain chemistry. Cognitive liberty concerns the ethics and legality of safeguarding one’s own thought processes, and by necessity, one’s electrochemical brain states. The individual, not corporate or government interests, should have sole jurisdiction over the control and/or modulation of his or her brain states and mental processes.”

In the same article, she also opines that CL “is the necessary substrate for just about every other freedom” [28]. This conception of CL as a prerequisite for other freedoms or rights is also defended by Ienca and Andorno [9].

Despite the more generic definition of CL offered by Boire, the most widespread way of understanding this principle builds on Sententia’s allusion to an individual’s control of the “modulation of his or her brain states and mental processes”—although, in our opinion, it would be perhaps more appropriate

¹² Original text (in Spanish): “*ser libre* significa poseer la capacidad para hacer lo que queremos, deseamos o escogemos hacer, y [también] que no haya [coerción o restricción] que nos impida que ejerzamos dicha capacidad” [21]. Translated to English by us.

to talk about *brain processes* and *mental states*. Following this line of thought, Bublitz [29] defines CL as “the right to alter one’s mental states *with the help of neurotools* as well as to refuse to do so” (our italics). This approach of referring specifically to the role of neurotechnologies, which we will employ here as our working definition, also implies that there are two simultaneous rights within CL: the *positive right* to use neurotechnologies and the *negative right* to be protected against coercive or non-consensual uses of these technologies [9, 29, 30].

In summary, the working definitions of the two neurorights concerning personal autonomy would be as follows:

- *Free will*: the right to “have ultimate control over [one’s] own decision making, without unknown manipulation from external neurotechnologies” [10].
- *Cognitive liberty*: “the right to alter one’s mental states with the help of neurotools as well as to refuse to do so” [29].

Finally, are the neurorights to FW and CL dependent on each other, complementary, or mutually exclusive? While the former intends to guarantee ultimate control of one’s decision making, the latter intends to protect one’s control of manipulating (or refusing to manipulate) one’s brain and, consequently, mental states. Since both neurorights are proposed from a practical point of view, they should be interpreted in the context of human actions. In turn, the right to FW openly mentions decision-making, which is unquestionably part of acting; according to the definitions given above, CL goes beyond action—or “more deeply” than action—but, in the context of personal autonomy, which has been defined as “the capacity to decide for oneself and pursue a course of action in one’s life” [14], this right is also intrinsically related to human action. In the following section, we will expose the challenges emerging from the mainstream understanding of human action, and more precisely, decision making, that are relevant to the neurorights concerning personal autonomy. We believe that a holistic understanding of human action may be more adequate to re-formulate these neurorights.

Decision, Intention, Action Realization: The Triple Dimension of Action

Another crucial conceptual clarification, with great relevance for personal autonomy, has to do with the understanding of human actions. For clarity, we are circumscribing our analyses here to voluntary actions, that is, actions that are purposefully performed to achieve a goal. To the best of our knowledge, no gold-standard action theory is referred to by the proponents of neurorights. One of the more successful and clear theories in the contemporary fields of cognitive sciences and philosophy of mind is Elisabeth Pacherie’s, whose approach consists in meticulously analyzing the different components of actions, namely intentions, deliberations, decisions, and actions themselves—that is, the ultimate realization of actions. As explained in Section [The Neuroright to Cognitive Liberty: How Does it Relate to Free Will?](#), decision making appears as a crucial element in the neurorights involving personal autonomy.

In action theory, decision is sometimes understood as closer to deliberation and, at other times, as closer to action realization. This is relevant to the argument presented in this article since it hinders the interpretation of cases in which an external neurotechnological intervention may be able to modify the final meaning of an action. When, within the framework of decision theory, an experiment is carried out where it is necessary to choose between two options, it is usually understood that there is deliberation before the choice. “Decision” is understood here as a trigger for action. However, it is worth pondering the following question: Is *decision* located on the edge of deliberation or on the edge of action?

In the case of Pacherie’s action theory, when this word appears in sentences like “The upstream dynamics of F-intentions [future-directed intentions]—the dynamics of decision-making that leads to the formation of an intention...” [31], this notion is closer to deliberation than action and is also prior to intention. Intentions are formed by decisions. In other places, expressions like “decision to start acting” and “decisions to act” [32] can be found in which decision seems to be closely related to action. In this case, intentions are thus prior to decisions. Although it is difficult to find a precise definition of intentions,

these are usually understood as mental states with two characteristics: being accessible to consciousness and being related to the subsequent action. A distinction has usually been made between prospective (a.k.a. distal) and immediate (a.k.a. proximal) intentions. These differ in their distance from the start of the action: “[P]rospective intentions necessarily occur before immediate intention and before action itself, and often long before them. In contrast, immediate intentions occur in the specific context of the action itself” [33]. In the case of prospective intentions, these “involve making a number of decisions. The intention is first formed when one reaches a decision about what to do. Once the intention is formed, one must still typically make a number of decisions about how to implement the chosen goal” [33]. In the first part of this description, it is understood that decision precedes intention, but in the second part, it seems that it is after the intention that a decision is made to implement it.

In any case, actions are constituted by intentions, decisions, and action realizations (or actions *per se*), all of them preceded by deliberations—the initial pondering concerning the action. The lack of specificity of decisions, understood on some occasions as prior to intention—and therefore as associated with deliberation—and on other occasions as linked to the beginning of the action, is relevant in neurorights, where autonomy in *decision making* is intended to be safeguarded. On the other hand, the right to CL proposes the capacity to intervene in one’s brain states, without referencing the possible effect on action. This ambiguity could be overcome by considering actions more holistically, instead of analyzing their components as if they were independent entities that link to each other to compose a chain of events. Hence, we propose to conceive action as a unified process, where intention and decision are defining elements of actions but should not be broken off from action realization. Intention, decision, and action realization constitute the *triple dimension of action*: they are different aspects of actions but not independent elements that can exist prior to action realization. In other

words, the three of them are not sequential but synchronic.¹³ One typical counterexample to this could be the following:

The calorie gain:

Bill has *decided* to go out running at 7 am. Bill wakes up, puts on his running clothes, and opens the door. However, it is heavily raining, so he turns around and has a beautiful breakfast.

Our response to this apparent counterexample is that Bill has *deliberated* to go running, but he has not *decided* to go running. As a matter of fact, he

¹³ Here is an analogy to play with: thinking of decision as a kind of “decisive momentum” of action. Only with reference to action can we truly speak of intention and decision-making. Only in action do we know not just that we *were*, but also that we *are* faced with a decision, and not a mere desire or purpose—no matter how elaborate it may be. Decisive momentum—i.e. decision—cannot be changed in itself; what can be modified is the process that results in the outcome of action.

The notion of momentum that we suggest does not allude to the word “moment” in either its logical or its phenomenological or its chronological sense, but is analogous to the momentum studied by Newtonian mechanics. In this discipline, both linear momentum p and angular momentum L are magnitudes that express the quantity of motion of a body, which remains constant under certain conditions of symmetry (translational or rotational) in a system. The angular momentum L of a rotating body relates its linear momentum ($p = \text{mass} \times \text{velocity}$) to the position vector r with respect to the point of origin O . What interests us here is not so much the formula with which L is expressed ($L = r \times p$), but rather to appreciate that p exists with respect to a point O , and also that L cannot be understood without an origin O to which p refers, nor without a moving body with p .

Now, here is the analogy. Just as the existence of the angular momentum L cannot be understood without p , the existence of a decision—decisive momentum—cannot be understood without an action. Moreover, just as it is not correct to say that p is part of L , nor that L is prior to p , neither is action part of decisive momentum, nor is decisive momentum prior to action. Quite simply, linear momentum p is accompanied by angular momentum L , and L is accompanied by p ; similarly, action is accompanied by decisive momentum, and decisive momentum accompanies action. Furthermore, just as a point O is needed to guide L , a goal is also necessary to guide decisive momentum. The point O thus indicates purpose or intention.

has decided to have an excellent breakfast, which is the real action that he has realized. Please note that deliberation is out of the triple dimension of action, and this solves the ambiguity concerning decisions in mainstream action theories: deliberation is distal concerning action, which is synchronically constituted by intention, decision, and action realization. Each of these dimensions can be analyzed independently to understand an action better, but they should not be understood as independent links of a chain.

Toward a Neuroright to Personal Autonomy

We have finished Section [Decision, Intention, Action Realization: The Triple Dimension of Action](#) by outlining how to interpret the triple dimension of action in response to the case of Bill exchanging his exercise for a delicious breakfast. This is (obviously) not a case of neurotechnological intervention, so it is irrelevant to the neurorights debate. Now, we will consider two practical cases that could indeed be real in the case of neurotechnologies. Mental experiments are usually confusing when they refer to fictional scenarios, such as desire or thought insertions, so we are restricting our case examples to situations that can already happen or could be real in the near future. Let us see the first case, raised by other researchers:

The twofold injury:

“A paralysed man participates in a clinical trial of a brain–computer interface (BCI). A computer connected to a chip in his brain is trained to interpret the neural activity resulting from his mental rehearsals of an action. The computer generates commands that move a robotic arm. One day, the man feels frustrated with the experimental team. Later, his robotic hand crushes a cup after taking it from one of the research assistants, and hurts the assistant. Apologizing for what he says must have been a malfunction of the device, he wonders whether his frustration with the team played a part” [13].

Has the neuroright to FW been violated in this case? Remember that, according to the NeuroRights Foundation [10], FW is the right to “have ultimate control over [one’s] own decision making, without unknown manipulation from external

neurotechnologies.” Besides the problematic inclusion of the concept of ultimate control, already mentioned in Section [“The Longstanding Problem with Free Will”](#), the critical question here is whether the BCI has externally manipulated the man’s *decision-making*. The problem that we highlighted in Section [“Decision, Intention, Action Realization: The Triple Dimension of Action”](#) arises: What is, specifically, the *decision* allegedly manipulated? Is it the decision that leads to the man’s frustration with the experimental team, or is it the decision to crush the cup out of his frustration? In the first assumption, it is a distal decision, prior to intention and far from action realization; in the second assumption, it is a proximal decision, after intention and close to action realization. As can be seen, a mainstream interpretation of action, in which it is composed of a chain of events of decision, intention, and realization, makes the ethical and normative interpretation of the case ambiguous. On the contrary, if we consider action holistically and synchronously, action is a unified process composed of three explanatory but *not sequential* dimensions—decision, intention, and realization. As a result, the critical question is simplified and specified in the following: Has the BCI participation been relevant to the outcome of the action (i.e., the cup crushing)?

Let us see now the second case, of our authorship:

The football cheater:

Mike is a football player that faces a decisive game. Since he plays as strong safety, he constitutes the last line of defense to avoid the rival team from scoring. Lastly, he is not feeling physically fit but thinks he can balance it with extra motivation. Thus, he intervenes in his neural system—through transcranial magnetic stimulation (TMS), since drugs could be detected in anti-doping tests—to increase his aggressiveness temporarily, only during the game. He feels great, strong as ever. During the game, when a rival is trying to cross his defensive line, Mike goes all out for him with unnecessary strength and causes a fatal injury to the rival player. Mike wonders whether the TMS made a difference or if the injury would have occurred anyway.

Has Mike’s neuroright to FW been violated? That is, has Mike’s decision-making been externally manipulated because of TMS? What is Mike’s relevant decision, the distal decision that precedes his

increased aggressiveness or the proximal decision to move toward his rival after his boost of aggressiveness? Again, understanding action as a chain of events makes the ethical and normative interpretation of the case too complex; but, according to our holistic, non-sequential interpretation, the critical question becomes no less than whether the TMS has been relevant in the outcome of the action (i.e., the injury of Mike's rival).

How can CL be interpreted in these two cases? According to our working definition, CL is "the right to alter one's mental states with the help of neurotools as well as to refuse to do so" [29]. In *The twofold injury*, the man agrees to participate in the clinical trial to get a BCI. Also, in *The football cheater*, Mike voluntarily intervenes on his neural system through TMS. In both cases, the protagonists exercise their right to alter their mental states, that is, their right to CL. Given this, it seems possible to interpret the right to CL separately from the right to FW, as the right to CL proposes the capacity to intervene in one's mental states, without explicit reference to the possible effect on action. Thus, a strategy would be to reformulate the right to FW, considering the difficulties we explained in Section "The Longstanding Problem with Free Will", and keep CL as an independent right. However, altering one's mental states (or refusing to do so) can also be interpreted as an action; we propose thus a suitable alternative strategy: to disambiguate both rights under a single *neuroright to personal autonomy*. We believe that this parsimonious strategy has the advantage that it could give proposals to protect autonomy greater normative effectiveness without duplicating efforts.

In our original working definition, personal autonomy is "the capacity to decide for oneself and pursue a course of action in one's life" [14]. In order to propose a neuroright to personal autonomy, two necessary modifications should be made:

(1) Since this is a metaphysical definition of autonomy (as it is understood as a capacity), we will adopt—as mentioned in Section "The Neuroright to Cognitive Liberty: How Does it Relate to Free Will?"—a practical stance because FW and CL become elements of a practical autonomy, rather than ontological constituents of personal autonomy.

(2) Since this definition understands autonomy as deciding for oneself, but decision is—according to our interpretation suggested in Section "Decision, Intention, Action Realization: The Triple Dimension of Action"—one of three non-sequential dimensions of autonomy, we will understand autonomy as a self-directed process that synchronously includes decision, intention, and action realization.

Following this approach, let us reformulate the original definition of a neuroright to

Personal autonomy:

the right to self-direct and pursue a course of action in one's life.

Thus, an action (including decision, intention, and realization) is considered *autonomous* as long as it is an outcome of this self-directed process without unauthorized or coercive intervention from neuro-technological sources. According to this interpretation (which encompasses the objectives of both FW and CL), in order for the neuroright to personal autonomy to be guaranteed in the two proposed cases, it would be necessary that:

- The protagonists requested or authorized the neuro-technological intervention (which happens), AND
- They were fully informed about the behavioral consequences that such intervention could entail. In this way, any unexpected outcome of their actions (i.e., cup crushing, injury of Mike's rival) could be attributed to a shared, anticipated, and permitted responsibility between the biological and the technological/mechanical parts.

Only by accomplishing both conditions could Mike and the paralyzed man be considered fully self-directing and pursuing the course of their actions. In other words, only by accomplishing both conditions could the neuroright to personal autonomy be preserved. According to this framework, decisions cannot be changed, as they are inseparable from actions. Changing the outcome of a decision can only be achieved by manipulating the self-directed process of action in which it is included as an explanatory component. This process may lead to altering, or not, one's mental states, (i.e. the

man participating in the BCI study, Mike undergoing TMS,) or it may be a subsequent process that is a consequence of the outcome about whether to alter one's mental states (i.e., the movement that leads to the cup crushing or the injury of Mike's rival).

It is important to emphasize that the "right to personal autonomy" is not a new initiative. However, there is no clear definition of what this right implies and what protections it covers, which may be quite different in each country. For example, in the U.S., as the Legal Information Institute [34] highlights, "The Supreme Court does not use the phrase 'personal autonomy' very often. Unlike privacy, it *is not a fundamental right*. As such, it is *still a very limited concept* regarding its impact on legal jurisprudence" (italics ours). Moreover, according to this Institute:

"In *Planned Parenthood v. Casey* (1992),¹⁴ the Court emphasized the impact that *Roe v. Wade* (1973)¹⁵ had on the importance of personal autonomy, especially with regard to reproductive rights. The *Casey* Court wrote, '[I]f *Roe* is seen as stating a rule of personal autonomy... [then the Supreme Court's] post-*Roe* decisions accord with *Roe*'s view that a State's interest in the protection of life falls short of justifying any plenary override of individual liberty claims... [N]o erosion of principle going to liberty or personal autonomy has left *Roe*'s central holding a doctrinal remnant.'

In *Washington v. Glucksberg* (1997)¹⁶ however, the Court appeared to oppose the concept that personal autonomy creates personal protections for individuals. 'And although *Casey* recognized that many of the rights and liberties protected by the Due Process Clause sound in personal autonomy, it does not follow that any and all important, intimate, and personal decisions are so protected. *Casey* did not suggest otherwise.' Evidently, the significance of a right to personal autonomy is far from certain." [34]

In Europe, if we look at the case of Spain, its Law 39/2006 for the Promotion of Personal Autonomy and Care for People in Situations of Dependency defines autonomy as "the ability to control, face, and make,

on one's own initiative, personal decisions about how to live in accordance with one's own norms and preferences as well as to develop the basic activities of daily life."¹⁷

As seen from these two brief examples, the right to personal autonomy can refer to aspects as divergent as reproductive rights or the promotion of the development of daily activities in dependent people. Therefore, there is room for expanding this right, including the new challenges neurotechnology brings. In any case, our proposal on personal autonomy is restricted to neurorights, that is, in cases involving a neurotechnological intervention. As such, it should be cautiously translated into other debates. Also, it should be noted that, although FW and CL (discussed in this work) are negative rights,¹⁸ other protections, such as that of dependent people, refer to positive rights, as they require active intervention by the State. Thus, what protections are guaranteed by personal autonomy, and whether or not they require intervention by the State (positive v. negative right) will require further discussion.

Concluding Remarks and Recommendations: Protecting Thought and Autonomy

The recent frenetic progress of neurotechnology earnestly invites us to reflect on its possible consequences, not only in the form of positive outcomes for health purposes but also in terms of potential risks to people's freedoms and liberties. Within this general context, we value the view that reforming the current human rights framework may be a suitable way to address these risks. The analyses and proposals we have made throughout this article aim to contribute to the conceptual and normative development of these neurorights, mainly those related to personal autonomy: FW and CL. We must insist that we have not tried to provide definitive solutions but to explore

¹⁷ Original text (in Spanish): "la capacidad de controlar, afrontar y tomar, por propia iniciativa, decisiones personales acerca de cómo vivir de acuerdo con las normas y preferencias propias así como de desarrollar las actividades básicas de la vida diaria." <https://www.boe.es/buscar/doc.php?id=BOE-A-2006-21990>. Accessed 18 Mar 2023. Translated to English by us.

¹⁸ To be more exact, according to our working definition of CL, this right has a twofold interpretation as a negative *and* positive right.

¹⁴ <https://www.law.cornell.edu/supct/html/91-744.ZS.html>

¹⁵ <https://supreme.justia.com/cases/federal/us/410/113/>

¹⁶ <https://www.law.cornell.edu/supremecourt/text/521/702>

some clues that may allow the discussion to move forward.

We recommend focusing on an umbrella right that can encompass the protections for personal autonomy provided by both FW and CL, and which can also give rise to other legislative enterprises where these rights' scope can be clearly defined. This umbrella right, i.e., personal autonomy, would protect the *externalization* of thought and the course of action, and it also would complement very well with a neuroright specifically focused on protecting the *internalization* of thought. Freedom of thought could be this umbrella right. However, it would probably be necessary to update it by expanding its scope of protection (currently focused primarily on freedom of religion and conscience) to meet the new challenges posed by neurotechnology. We agree with Ienca [8] in his opinion that “[a]n evolutionary interpretation of this right should focus on the protection not only of externalizations of thought but thought itself.” This importance of protecting thought itself should guide the expansion to which we refer; article 18 of the UDHR being an excellent niche to do so. This article reads as follows: “Everyone has the right to freedom of thought, conscience and religion; this right includes freedom to change [their] religion or belief, and freedom, either alone or in community with others and in public or private, to manifest [their] religion or belief in teaching, practice, worship and observance” [17]. As can be seen, this article is devoted to freedom of thought, but focuses on its externalizations, so there is conceptual and normative space also to include cognitive, internal processes.

Even though in philosophy, from a historical stance, FW has been considered in ontological terms (i.e., as an intrinsic property of the human being), the NeuroRights Foundation's proposal [10] transfers it from the metaphysical to the practical realm. Hence, both FW and CL become elements of a practical autonomy, rather than ontological constituents of personal autonomy. Should neurorights be considered and formulated in practical or in ontological terms? The detailed answer to this question is beyond the scope of this manuscript, but we would like to pose this concern explicitly, since many articles included in the UDHR point to the ontological nature of rights—rights that are owned by being a human person—while others emerge from the practical exercise of being human. We have adopted this practical stance in this manuscript, but further discussions about this

fundamental issue should be encouraged. Interestingly, in a view that is compatible with our practical stance regarding personal autonomy, Bublitz [35] has recently described the practical, non-metaphysical interpretation of freedom of thought:

“‘Thought is free’ has further meanings. It is usually not a claim about freedom in a metaphysical sense (as in the free will debate) or about freedom in the mind-brain relationship (as in substance dualism), but about an empirical inaccessibility of thought that is twofold: Undisclosed thoughts are free because they cannot be known by others, except perhaps by God; thoughts are free because they are factually invincible as they cannot be changed or prevented by others against the thinkers' will.”

In sum, we believe that our proposal of combining FW and CL in a single neuroright to personal autonomy, complemented by a neuroright to freedom of thought, can serve as a valuable tool for sharply delineating the distinction between the externalizations and internalizations of thought and action. While the former would all fit into a unified concept of action—that encompasses decision, intention, and action realization as explanatory, non-sequential elements—the latter would be understood as the set of deliberative processes leading to action. Ultimately, personal autonomy and freedom of thought are still underdeveloped rights with great scope for discussion and eventual expansion in light of the new neurotechnological challenges.

Protecting freedom of thought and personal autonomy is to protect traits that define us as humans. It is not possible to understand human beings and our nature without understanding our wonderful capacity for abstract, complex, and long-term goal-oriented behavior. Any effort made for the sake of freedom is certainly worth it, as “the treasures that the earth holds and the sea conceals cannot compare to it.”

Acknowledgements JM and JB want to thank the Tatiana Foundation (<https://fundaciontatianapgb.org/>) for its valuable support of their research. JM, JB, and FG are grateful to the Mind-Brain Group at ICS and, especially, to Nathaniel Barrett, Ph.D., and also want to thank John Horden for his English language review of the manuscript.

Author Contributions All authors contributed equally to the manuscript and approved the submitted version.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data Availability This manuscript has no associated data.

Declarations

Ethics Approval This article does not contain any studies with human or animal subjects performed by either of the authors.

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Reuse of Material Section “The Longstanding Problem with Free Will” includes material presented by JM at the 2021 Annual Meeting of the International Neuroethics Society (<https://www.neuroethicssociety.org/2021-annual-meeting-abstracts#p53>) and the Proceedings of the 10th Conference of the Spanish Society for Logic and the Methodology and Philosophy of Science (<https://xslmfce.usal.es/index.php/proceedings/>). This material has been included with the permission of both publishers.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Swetnam, M., McBride, D., Herzfeld, C., Barnett, J., Schiller, K., Gallington, D., Retelle, J., Siegrist, D., Buss, J., and J. Giordano. 2013. *Neurotechnology Futures Study*. Arlington: Potomac Institute for Policy Studies. <https://www.potomacinstitute.org/images/studies/NTFSReport-FINAL.pdf>. Accessed 18 Mar 2023.
- Steinert, S., C. Bublitz, R. Jox, and O. Friedrich. 2019. Doing things with thoughts: Brain-computer interfaces and disembodied agency. *Philosophy & Technology* 32: 457–482. <https://doi.org/10.1007/s13347-018-0308-4>.
- Rueda, J. 2021. Socrates in the fMRI scanner: The neuro-foundations of morality and the challenge to ethics. *Cambridge Quarterly of Healthcare Ethics* 30 (4): 604–612. <https://doi.org/10.1017/S0963180121000074>.
- Farina, M., and A. Lavazza. 2022. Incorporation, transparency and cognitive extension: Why the distinction between embedded and extended might be more important to ethics than to metaphysics. *Philosophy & Technology* 35: 10. <https://doi.org/10.1007/s13347-022-00508-4>.
- Ricci, G. 2020. Pharmacological human enhancement: An overview of the looming bioethical and regulatory challenges. *Frontiers in Psychiatry* 11: 53. <https://doi.org/10.3389/fpsy.2020.00053>.
- Kubaneck, J., Brown, J., Ye, P., Pauly, K.B., Moore, T., and W. Newsome. 2020. Remote, brain region-specific control of choice behavior with ultrasonic waves. *Science Advances* 6 (21): eaaz4193. <https://advances.sciencemag.org/content/6/21/eaaz4193>.
- Han, W., L.A. Tellez, M.R. Rangel Jr., S.C. Motta, X. Zhang, I.O. Perez, N.S. Canteras, S.J. Shammah-Lagnado, A.N. van den Pol, and I.E. de Araujo. 2017. Integrated control of predatory hunting by the central nucleus of the amygdala. *Cell* 168 (1–2): 311–324.e18. <https://doi.org/10.1016/j.cell.2016.12.027>.
- Ienca, M. 2021. On neurorights. *Frontiers in Human Neuroscience* 15: 701258. <https://doi.org/10.3389/fnhum.2021.701258>.
- Ienca, M., and R. Andorno. 2017. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy* 13: 5. <https://doi.org/10.1186/s40504-017-0050-1>.
- NeuroRights Foundation (n.d.). *Mission*. NeuroRights Foundation. <https://neurorightsfoundation.org/mission>. Accessed 18 Mar 2023.
- Yuste, R., Genser, J., and S. Herrmann. 2021. It’s time for neuro-rights. *Horizons* 18: 154–164. <https://www.cirsd.org/en/horizons/horizons-winter-2021-issue-no-18/its-time-for-neuro-rights>.
- Goering, S., Klein, E., Specker Sullivan, L., Wexler, A., Agüera y Arcas, B., Bi, G., Carmena, J. M., Fins, J.J., Friesen, P., Gallant, J., Huggins, J.E., Kellmeyer, P., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., Teicher, ..., and R. Yuste 2021. Recommendations for responsible development and application of neurotechnologies. *Neuroethics* 14: 365–386. <https://doi.org/10.1007/s12152-021-09468-6>.
- Yuste, R., Goering, S., Agüera y Arcas, B., Bi, G., Carmena, J. M., Carter, A., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., ..., and J. Wolpaw. 2017. Four ethical priorities for neurotechnologies and AI. *Nature* 551(7679): 159–163. <https://doi.org/10.1038/551159a>.
- Dryden, J. 2021. Autonomy. In *Internet encyclopedia of philosophy*, eds. J. Fieser and B. Dowden. <https://iep.utm.edu/autonomy/>. Accessed 18 Mar 2023.
- Bublitz, J.C., and R. Merkel. 2014. Crimes against minds: On mental manipulations, harms and a human right to mental self-determination. *Criminal Law and Philosophy* 8: 51–77. <https://doi.org/10.1007/s11572-012-9172-y>.
- Bublitz, J.C. 2016. Moral enhancement and mental freedom. *Journal of Applied Philosophy* 33: 88–106. <https://doi.org/10.1111/japp.12108>.
- United Nations General Assembly. 1948. *Universal Declaration of Human Rights*. United Nations. <https://www>.

- [un.org/en/about-us/universal-declaration-of-human-rights](https://www.un.org/en/about-us/universal-declaration-of-human-rights). Accessed 18 Mar 2023.
18. Muñoz, J.M. 2019. Chile — right to free will needs definition. *Nature* 574 (7780): 34. <https://doi.org/10.1038/d41586-019-03295-9>.
 19. Herrera-Ferrá, K., J.M. Muñoz, H. Nicolini, G. Saruwatari, and V.M. Martínez Bullé. 2022. Contextual and cultural perspectives on neurorights: Reflections toward an international consensus. *AJOB Neuroscience*: Online first. <https://doi.org/10.1080/21507740.2022.2048722>.
 20. Borbón, D.O., and L. Borbón. 2021. A critical perspective on NeuroRights: Comments regarding ethics and law. *Frontiers in Human Neuroscience* 15: 703121. <https://doi.org/10.3389/fnhum.2021.703121>.
 21. Muñoz, J.M. 2012. Hacia una sistematización de la relación entre determinismo y libertad. *Daimon* 56: 5–19. <https://revistas.um.es/daimon/article/view/141761>.
 22. Wolf, S. 2005. Freedom within reason. In *Personal autonomy*, ed. J.S. Taylor, 258–274. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511614194.012>.
 23. O'Connor, T., and C. Franklin. 2022. Free will. In *Stanford encyclopedia of philosophy*, eds. E.N. Zalta and U. Nodelman. Stanford University. <https://plato.stanford.edu/entries/freewill/>. Accessed 18 Mar 2023.
 24. Chernyak, N., C. Kang, and T. Kushnir. 2019. The cultural roots of free will beliefs: How Singaporean and U.S. children judge and explain possibilities for action in interpersonal contexts. *Developmental Psychology* 55 (4): 866–876. <https://doi.org/10.1037/dev0000670>.
 25. Berniūnas, R., A. Beinorius, V. Dranseika, V. Silius, and P. Rimkevičius. 2021. The weirdness of belief in free will. *Consciousness and Cognition* 87: 103054. <https://doi.org/10.1016/j.concog.2020.103054>.
 26. Wisniewski, D., R. Deutschländer, and J.-D. Haynes. 2019. Free will beliefs are better predicted by dualism than determinism beliefs across different cultures. *PLOS One* 14 (9): e0221617. <https://doi.org/10.1371/journal.pone.0221617>.
 27. Boire, R. G. 2000. On cognitive liberty (Part 2). *Journal of Cognitive Liberties* 2 (1), 7–20. <https://www.cognitive-liberty.org/ccle1/2jcl/2JCL7.htm>. Accessed 18 Mar 2023.
 28. Sententia, W. 2004. Neuroethical considerations: Cognitive liberty and converging technologies for improving human cognition. *Annals of the New York Academy of Sciences* 1013 (1): 221–228. <https://doi.org/10.1196/annals.1305.014>.
 29. Bublitz, J. C. 2013. My mind is mine!? Cognitive liberty as a legal concept. In *Cognitive enhancement*, eds. E. Hildt and A.G. Franke, 233–264. Springer. <https://doi.org/10.1007/978-94-007-6253-4>.
 30. Farahany, N.A. 2023. *The battle for your brain*. New York: St. Martin's Press.
 31. Pacherie, E. 2006. Towards a dynamic theory of intentions. In *Does consciousness cause behavior? An investigation of the nature of volition*, eds. S. Pockett, W.P. Banks, and S. Gallagher, 145–167. Cambridge: MIT Press. <https://doi.org/10.7551/mitpress/9780262162371.003.0009>.
 32. Pacherie, E. 2008. The phenomenology of action: A conceptual framework. *Cognition* 107 (1): 179–217. <https://doi.org/10.1016/j.cognition.2007.09.003>.
 33. Pacherie, E., and P. Haggard. 2010. What are intentions? In *Conscious will and responsibility: A tribute to Benjamin Libet*, eds. W. Sinnott-Armstrong and L. Nadel, 70–84. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195381641.003.0008>.
 34. Legal Information Institute (n.d.). *Personal autonomy*. Cornell Law School. https://www.law.cornell.edu/wex/personal_autonomy. Accessed 18 Mar 2023.
 35. Bublitz, J. C. 2022. Freedom of thought. In *Encyclopedia of the philosophy of law and social philosophy*, eds. M. Sellers and S. Kirste. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-6730-0_1066-1.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.