



Review article

Can neuroscience enlighten the philosophical debate about free will?

Claire Delnatte^a, Emmanuel Roze^{b,c}, Pierre Pouget^b, Cécile Galléa^b, Quentin Welniarz^{b,*}^a Education nationale, académie de Paris, France^b Sorbonne Université, Faculté de Médecine, INSERM U 1127, CNRS UMR 7225, Paris Brain Institute Institut du Cerveau, F-75013, Paris, France^c Assistance Publique – Hôpitaux de Paris, Hôpital Pitié-Salpêtrière, Département de Neurologie, Paris, France

ARTICLE INFO

Keywords:

Free will
 Determinism
 Compatibilism
 Libertarian
 Free choice
 Decision-making
 Mental causation
 Physicalism
 Conscious intention

A B S T R A C T

Free will has been at the heart of philosophical and scientific discussions for many years. However, recent advances in neuroscience have been perceived as a threat to the commonsense notion of free will as they challenge two core requirements for actions to be free. The first is the notion of determinism and free will, i.e., decisions and actions must not be entirely determined by antecedent causes. The second is the notion of mental causation, i.e., our mental state must have causal effects in the physical world, in other words, actions are caused by conscious intention.

We present the classical philosophical positions related to determinism and mental causation, and discuss how neuroscience could shed a new light on the philosophical debate based on recent experimental findings. Overall, we conclude that the current evidence is insufficient to undermine free will.

1. Introduction

Free will is tightly linked to the notion of volition and is a key feature of human nature (Roskies, 2010); we exert our free will through voluntary actions and decisions. Although the definition of free will is a matter of debate in neuroscience and philosophy, our commonsense experience of free will involves two core elements (Balaguer, 2009). First, our decisions and actions must not be completely determined by antecedent causes beyond our control. When facing a situation with different alternatives (for instance, when deciding which gift we should choose for a birthday or what meal we should have for lunch), an intuitive way to consider that we have chosen freely is that we could have done otherwise, all other conditions remaining the same. By contrast, if our actions were completely determined by antecedent causes, we could not have done otherwise, and hence we could not have chosen freely. This notion will be later referred to as the problem of free will and determinism (Honderich, 1990; Kane, 1998; Pereboom et al., 2001). Second, from our everyday experience it would seem that our mental state, and in particular our conscious intentions, can cause physical events. For example, I want to grab this glass of water because I am thirsty, and my conscious intention would thus cause my arm to move toward the glass. How a mental event can cause a physical event has been a longstanding question in philosophy and this second notion is usually referred to as the problem of free will and mental causation

(Walter and Heckmann, 2003).

These two problems had been a subject of debate in philosophy of mind for a long time before neuroscientists started to take position. Over the last decades, advances in neuroscience have expanded the possibilities for experimental investigations of these longstanding philosophical questions. In this review, we first discuss free will and determinism and then go onto discuss free will and mental causation. For each part, we start by presenting a comprehensive review of the classical philosophical positions of free will. We then consider how neuroscience could shed a new light on these notions by reviewing recent experimental data. Finally, we discuss the results from neuroscience, their limitations, and consider their contribution to the philosophical debate.

2. Free will and determinism

2.1. Philosophical questions

For our actions to be free, they must not be entirely determined by antecedent causes beyond our control. Although there is considerable debate about this definition, an intuitive way is to consider our actions free when we could have acted otherwise, all other conditions remaining the same (Dennett, 2015; Frankfurt, 1969; Kane, 1998; Pereboom et al., 2001). When we choose between two alternatives (for instance, staying at home to read a book or going for a walk), we may have reasons to

* Corresponding author. Institut du cerveau, Hôpital Pitié Salpêtrière, 47boulevard de l'Hôpital, 75013, Paris, France.

E-mail address: quentin.welniarz@gmail.com (Q. Welniarz).

prefer one to another, but we feel that our choice is free because choosing one option rather than the other is not predetermined. We may eventually decide to go for a walk, but equally could have chosen to stay home to read a book instead. This is in stark contrast to the physical world, at least at a specific scale. For example, the physical laws that govern the movements of particles entail a form of necessity. When we heat water over 100 °C, it starts boiling because it is determined by physical laws, and it could not happen otherwise.

Determinism can be defined as follows: every physical event is causally necessitated by prior events together with the causal laws of physics (Balaguer, 2009; Kane, 1998; Pereboom et al., 2001; Roskies, 2006; Searle, 2001). If the universe is indeed deterministic, we lack free will because everything, including our actions and decisions, is entirely determined by a prior state of the universe together with physical laws (Kane, 1998; Pereboom et al., 2001; Roskies, 2006). Accordingly, in the same way that water necessarily boils when it is heated over 100 °C, we cannot act otherwise than we do. Contrasted philosophical positions have emerged regarding this issue of free will and determinism (Roskies, 2006, 2010, 2012; Shadlen and Roskies, 2012). Compatibilists believe that determinism is true, but that free will is compatible with determinism (Bok, 1998; Frankfurt, 1971; Hume, 1960). For instance, according to Hobbes, a person is free when they act in accordance with their will and desires (Hobbes, 1962). Because will and desires can themselves be determined by antecedent causes, this definition of freedom is compatible with determinism. Thus, people are free unless they cannot act according to their will, for instance when they are under constraint, coercion, or when they suffer from compulsions or addictions. From the compatibilist's view, the ability to do otherwise, that is central to free will, can be interpreted as follows: you could have done otherwise if you had willed otherwise. By contrast, incompatibilists believe that determinism precludes freedom, because they consider that free will entails the ability to do otherwise *all other conditions remaining the same*. This definition of free will is clearly incompatible with determinism, because under the exact same circumstances, there is only one possible outcome according to the definition of determinism. This leads to two distinct philosophical trends. On the one hand, hard incompatibilists accept determinism and conclude that we do not have free will (Pereboom et al., 2001). On the other hand, libertarians believe that we have free will because some events are indeterministic (Kane, 1998; O'Connor, 1972). Indeed, following the discovery of quantum physics, the introduction of probabilistic, as opposed to deterministic, laws opened the possibility for indeterministic events in the universe (Conway and Kochen, 2006; Searle, 2001). However, indeterminism is also full of inherent problems: it introduces the notion of randomness in our behaviors, and consequently would seem to be incompatible with the definition of free will (Balaguer, 2014; Ebert and Wegner, 2011; Kane, 1998; Pereboom et al., 2001; Roskies, 2006; Searle, 2001; Shadlen and Roskies, 2012). Libertarians argue that even if probabilistic laws underlie our decisions, it does not preclude the possibility of an agent causing and controlling their actions. Probabilistic laws make our actions more or less probable, while leaving the eventual choice open (O'Connor, 2022). Overall, the philosophical debate regarding free will and determinism is not yet settled.

2.2. Inputs from neuroscience

To examine how neuroscience can contribute to the discussion about free will and determinism, we should examine how free will can be operationalized for experimental studies. Freely willed actions are typically defined as internally-generated or self-initiated, in contrast to externally-triggered or stimulus-driven actions (Fried et al., 2017; Haggard, 2008; Roskies, 2010; Seghezzi et al., 2019; Seghezzi and Haggard, 2022). In front of a traffic signal, our behavior is externally triggered: we stop when the light is red, and we go when it turns green. By contrast, our decision to go for a walk rather than staying at home to read a book is to all extents and purposes internally driven: it could

depend on our belief that the weather forecast predicts a sunny afternoon or on our desire to stretch our legs after a long day at work. This view is related to the conception of free will as a form of decision-making. Free actions could be seen as a succession of internally-generated decisions regarding what goal to pursue or what action to perform (the “what” component), when to do it (the “when” component), and whether to do it (the “whether” component) (Brass and Haggard, 2008). When transposed to the laboratory, this results in the so-called “free-choice” paradigm. Subjects are instructed to freely choose between different actions (the “what” component, for instance, between right- and left-hand movements); to decide the moment to initiate the action (the “when” component); or to decide whether to execute or to withhold their action (the “whether” component). These conditions can then be compared with cases in which the action is completely determined by external stimuli (stimulus-driven action), instructing the subjects what action to perform, and when or whether to perform it. The “free-choice versus stimulus-driven action” paradigm is a good tool to investigate changes in brain activity associated with the decision-making process and neural substrates of free choices, albeit with some limitations (Bode et al., 2014; Haggard, 2008, 2019).

Thus, characterizing the neural mechanisms underlying free choice should provide further insight into the question of determinism and free will (Roskies, 2006, 2010, 2012; Searle, 2001). If it turns out that the brain mechanisms underlying our free choices are entirely deterministic, this would favor determinism and consequently threaten our common-sense notion of free will. By contrast, if our decisions are due to indeterministic (random, stochastic) events in our brain, some libertarians could take this evidence as proof that we have free will.

2.3. Results from neuroscience

Non-invasive neuroimaging and neurophysiological exploration represent convenient and accurate approaches to characterize the properties of brain networks associated with a particular event. We will review the neuroscientific data about free will, here defined as “free choice”.

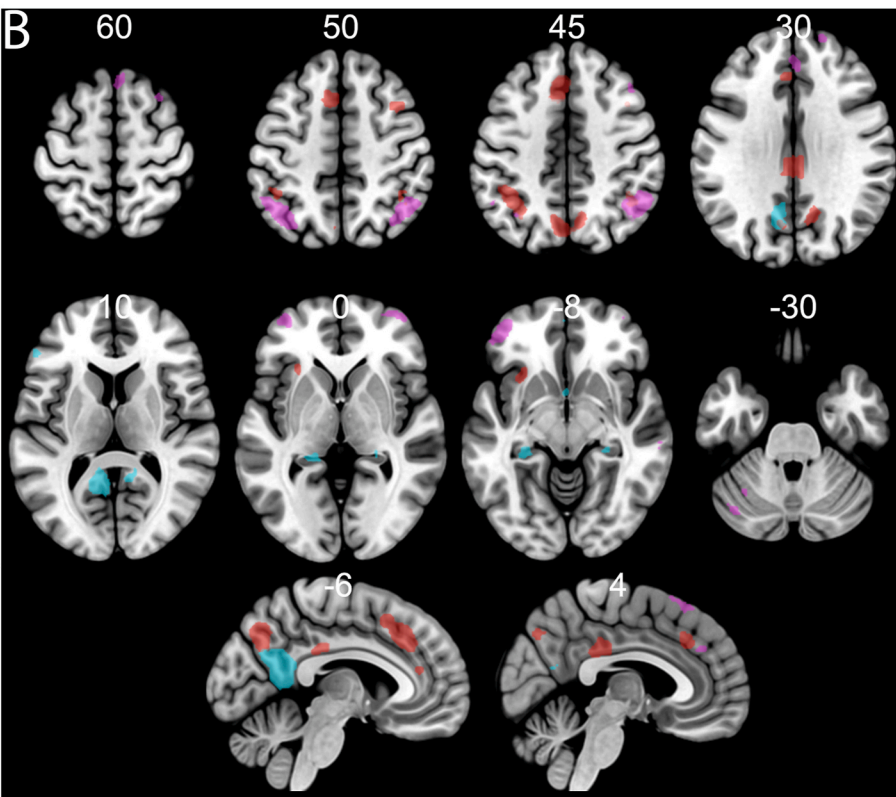
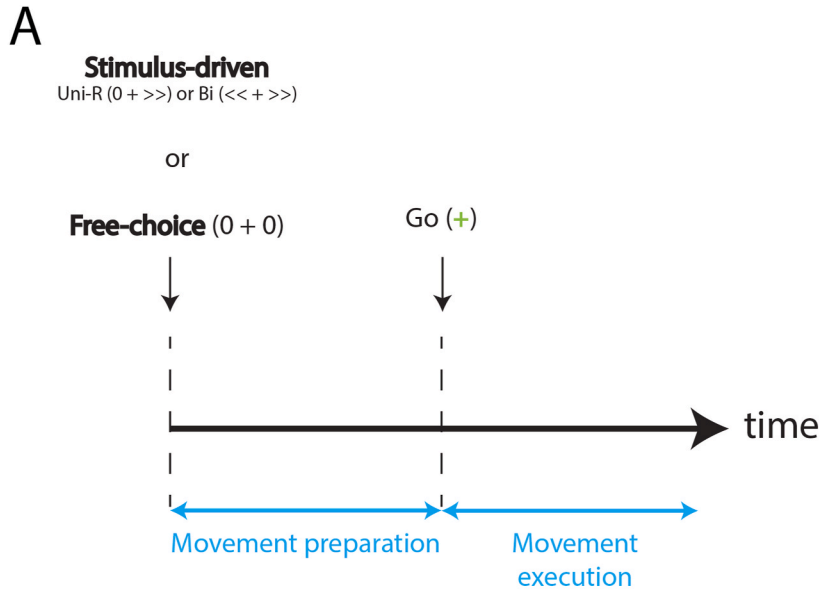
2.3.1. Neuroanatomy of free choice

2.3.1.1. fMRI experiments. The previously described “free choice” protocol has been repeatedly used in functional MRI (fMRI) studies. Trials in which the subjects can freely choose what, when or whether to perform a given action have been compared with trials in which the same action is completely determined by external cues. This comparison led to the identification of brain regions and networks that are more engaged in free-choice actions than in stimulus-driven actions (Haggard, 2008; Seghezzi et al., 2019; Seghezzi and Haggard, 2022; Zapparoli et al., 2017). These studies have recently been reviewed and were the subject of three meta-analyses (Seghezzi et al., 2019; Seghezzi and Haggard, 2022; Zapparoli et al., 2017). In one of the reports, the “what”, “when” and “whether” components were analyzed separately (Zapparoli et al., 2017): the “what” component was associated with greater activation in the cingulum, the frontal middle gyrus, the supramarginal gyrus, and the inferior triangular frontal gyrus; the “when” component with greater activation in the supplementary motor area (SMA), the middle frontal gyrus, and the inferior parietal lobule; and the “whether” component with greater activation in the cingulum, the inferior orbital frontal gyrus, the insula, and subcortical structures. These results were confirmed in a study that directly contrasted each individual component (“what”, “when” and “whether”) against the two others (Zapparoli et al., 2018). In the two other meta-analyses, the three components (“what”, “when”, and “whether”) were pooled. Self-generated actions, as compared with stimulus-driven actions, were associated with stronger activation in the middle cingulum, the middle and inferior frontal gyrus, the SMA and pre-supplementary motor area (preSMA), the anterior

insula, the superior and inferior parietal lobules, and the cerebellum (Seghezzi et al., 2019; Seghezzi and Haggard, 2022).

Intuitively, it would seem that we exert our free will at the moment of the deliberation phase that precedes the decision and the action. Movement execution and the underlying brain activity should be unaffected by whether the movement was freely chosen or stimulus-driven. In consequence, many of the previous fMRI studies addressing this

question did not use an event-related design, so that movement execution and preparation could not be disentangled in the analysis (Deiber et al., 1991; Filevich et al., 2013; Forstmann et al., 2006; Hoffstaedter et al., 2013; Lau et al., 2004; Mueller et al., 2007; van Eimeren et al., 2006). By contrast, other studies focused on the preparation phase rather than on movement execution (Ariani et al., 2015; Krieghoff, 2009; Zapparoli et al., 2018). We recently tackled this question in a free



Movement preparation

Movement execution

■ Free choice > stimulus-driven

■ Free choice > stimulus-driven

■ Stimulus-driven > free choice

Fig. 1. “Free choice” task: identification of the brain activity underlying free choice.

(A) Organization of a trial. During the delay period, the subject had to prepare the forthcoming movement (right-hand movement or bimanual movement) that was either stimulus-driven or freely chosen. In the stimulus-driven situation, a first cue indicated to the subject which movement to perform (right-hand or bimanual). In the free-choice condition, the decision was up to the participants. After the “Go” signal, the subject had to execute the prepared movement. (B) By comparing the trials in which the movement was freely chosen by contrast to stimulus-driven, we can identify the anatomical location brain areas more activated during freely chosen as compared to stimulus-driven movements during the preparation phase (red) and execution phase (purple). By contrast, some regions were more activated during stimulus-driven movements during the execution phase (light blue).

Uni-R: unimanual right-hand movement; Bi: bimanual movement.

Adapted from (Welniarz et al., 2021).

choice experiment, in which we analyzed movement preparation and execution separately (Welnarz et al., 2021). During the preparation phase, participants were asked to prepare either a right-hand or bimanual movement. The selection between these two possibilities was either freely chosen or stimulus-driven by an external cue. During the execution phase, the participants had to execute the prepared movement after a “Go” signal (Fig. 1A). Nonetheless, the temporal window we used in this fMRI analysis is questionable. Because free will lies in the ability to do otherwise, the deliberation phase should theoretically extend up to the “point of no return”, immediately preceding the movement’s onset. None of the aforementioned studies used such a temporal window to contrast internally-generated and externally-cued movements.

In our study, we confirmed that the preparation of freely-chosen as compared to stimulus-driven movements was associated with increased activity in several brain regions, including the preSMA, the middle cingulum, the anterior cingulate cortex, the inferior parietal cortex, the middle frontal gyrus, and the insula (Fig. 1B). Interestingly, we found that, compared with externally-cued movements, the execution of the freely-chosen movements was associated with different patterns of brain activity. We found increased fMRI signals in the inferior parietal cortices, the SMA, the cerebellum, and frontal lobe areas (Fig. 1B). Using connectivity analyses, we showed that freely-chosen movements involved increased connectivity between the inferior parietal cortex and the cerebellum. We suggest that the execution of freely-chosen movements is associated with functional interaction in a network that monitors the congruence between the intentional content of our actions and their outcomes.

2.3.1.2. Electrophysiology. Electrophysiology experiments have also contributed to the description of the brain regions underlying self-initiated actions. EEG recordings in healthy subjects have consistently identified a negative potential that precedes voluntary movements by approximately 1 s: the readiness potential (RP) (Dirnberger et al., 1998; Kornhuber and Deecke, 1965; Shibasaki and Hallett, 2006). The RP has long been considered as an important brain signal for the generation of voluntary movements, although this classical view has recently been challenged (see Section 2) (Schurger et al., 2021). The early component (approximately 1500–400 ms before movement onset) consists of a slow and gradual negativity which is generated by the preSMA, further indicating a role of this region in brain activation related to free choice. The role of the SMA in the generation of the early component of the RP was confirmed by intracranial recordings of single neurons (Fried et al., 2011). By contrast, the late component (400-0 ms prior to movement onset) is generated by the primary motor cortex (Haggard, 2008; Roskies, 2010; Shibasaki and Hallett, 2006; Yazawa et al., 2000). However, neurostimulation studies using transcranial magnetic stimulation (TMS) or transcranial direct cortical stimulation (tDCS) showed that stimulation of the primary motor cortex failed to influence hand selection in a free-choice task (Sohn et al., 2003; Thirugnanasambandam et al., 2019). Thus, modulating the excitability of the final effector of the motor pathway does not influence the voluntary generation of movements. In addition, electrophysiological recordings in non-human primates comparing free and externally-cued choices led to the identification of the parietal cortex as a key region for free choice (Cui and Andersen, 2007; Pesaran et al., 2008).

2.3.1.3. Lesion studies. Neuropsychology investigates brain functions by studying behavioral modifications induced by cerebral damage to specific regions. In this way we can identify the regions of the brain that are involved in a functional network underlying a specific cognitive function (Vaidya et al., 2019). The aforementioned fMRI and electrophysiological studies suggest that the prefrontal cortex (in particular the SMA and preSMA) and the parietal cortex are strongly involved in the generation of voluntary actions. Accordingly, lesions of these two regions are associated with either a decrease in self-initiated movements

or an excess of involuntary actions (Desmurget and Sirigu, 2012; Fried et al., 2017; Haggard, 2008; Hallett, 2007, 2016; Seghezzi and Haggard, 2022). Lesions of the preSMA have been associated with a decrease in self-initiated actions in non-human primates and humans (Kravnik et al., 2004; Laplane et al., 1977; Nachev et al., 2008; Passingham, 1987; Passingham et al., 2010; Potgieser et al., 2014; Thaler et al., 1995). By contrast, lesions of the SMA and parietal cortex can result in involuntary movements of the contralateral hand, a phenomenon known as “alien hand syndrome”. A unilateral lesion of the SMA results in involuntary movements in response to external cues: patients automatically reach for objects that are in the visual space contralateral to the lesion. Lesions of the parietal cortex result in purposeless movements without any intention, such as “hand levitation” or “hand wandering”, in the absence of an external cue (Assal et al., 2007; Desmurget and Sirigu, 2012; Fried et al., 2017; Hassan and Josephs, 2016). Conversely, psychiatric disorders manifesting an abnormal experience of free will, such as schizophrenia, are associated with altered activity of the parietal cortex. In the passivity phenomenon, self-generated actions and thoughts are experienced as being external. Passivity symptoms in schizophrenic patients and healthy individuals correlate with an increased activity in the parietal cortex and cerebellum (Blakemore et al., 2003; Schnell et al., 2008; Spence et al., 1997). This misattribution of internally-generated movements or thoughts in schizophrenic patients could originate from a failure to predict the sensory consequence of their own actions (Bansal et al., 2018).

In conclusion, our ability to choose freely does not rely on a single brain region, but rather on a widespread network of which the SMA, the anterior cingulate cortex, the posterior parietal cortex, and the cerebellum are key nodes (Haggard, 2008; Pesaran et al., 2008; Welnarz et al., 2021).

2.3.2. Neural mechanisms of free choice

The association of free will (as operationalized in neuroscience) with a specific brain activity is not surprising for the scientists and philosophers who accept ontological physicalism, particularly for those who think that mental properties supervene on physical properties. According to this view, the relation between mental and physical events goes beyond a mere covariation. Mental events are dependent on, and determined by, physical events in the brain. But these results give us no direct answer regarding the question of determinism and free will. We need to delve deeper into the neural mechanisms that underlie free choice to see whether these processes are deterministic or indeterministic.

Our current understanding of the neural mechanisms of decision-making largely relies on electrophysiological studies in non-human primates. The general framework was first set by the study of perceptual decision-making in monkeys, where the animal has to make a decision between different possibilities based on sensory evidence (Gold and Shadlen, 2007). Presented with a cloud of dots randomly moving on a screen, the monkey has to determine the direction of the net motion and then has to make an ocular saccade in the same direction to indicate its response. According to the “accumulation-to-bound” model, the different possible responses are represented by different neuronal populations in the lateral intraparietal area that is part of the posterior parietal cortex. The firing rate of these groups of neurons encodes the accumulation of sensory evidence until one of them reaches a threshold that triggers the response (Gold and Shadlen, 2007; Roskies, 2010). This model was then extended to more complex situations, and it can be assumed that the neurons’ firing rate can encode various aspects of the decision-making process such as the expected value of the outcome and its probability (Gold and Shadlen, 2007; Roskies, 2010; Shadlen and Roskies, 2012).

If we consider free will in the sense of self-generated actions, the decision should be triggered by internal factors rather than sensory evidence. The accumulation-to-bound model may be applied in the context of free choice, i.e., in the absence of external evidence that could

be used to determine the action (Bode et al., 2014; Brass et al., 2019; Murakami et al., 2014; Roskies, 2010; Schurger et al., 2016). In such cases, what kind of internal signal could be used to make the decision? Electrophysiological experiments in humans, non-human primates, and rodents indicate that when external evidence is lacking or ambiguous, the accumulation of stochastic neural noise makes the signal cross the decision threshold (Brass et al., 2019; Murakami et al., 2014; Roskies, 2010; Schurger et al., 2012, 2016). In this regard, free choice would rely, at least partially, on random neuronal noise, and may not be entirely deterministic. Libertarians would consider this as sufficient proof in favor of free will, whereas others would consider that randomness is as incompatible with free will as determinism (Shadlen and Roskies, 2012). Although it may be hard to consider that noise and free will are compatible, these scientific results suggest that there is no contradiction between stochastic neural activity and intentional action.

Another line of research suggests that in the absence of external evidence, free will is a result of the accumulation of contextual evidence rather than of neuronal noise. Contextual evidence could include the task instruction or the history of recent choices, even if this kind of information is usually not consciously monitored (Bode et al., 2014; Hwang et al., 2017; Mochizuki and Funahashi, 2014). In contrast with the previous consideration relating to random neuronal noise, the notion of contextual evidence suggests that the neural substrates of free choice are deterministic. According to incompatibilists, such a conclusion would preclude free will.

2.4. Can neuroscience enlighten the philosophical debate regarding free will and determinism?

The picture that emerges from the studies of the neuronal mechanisms underlying free choice leaves us with two options. On one hand, that the neuronal substrates of free choice are fundamentally indeterministic as they rely on random neuronal noise (Brass et al., 2019; Roskies, 2010; Schurger et al., 2012, 2016, 2021). On the other hand, that our so-called “free choice” could be the result of an accumulation of unconscious contextual evidence, a view that leans toward determinism (Bode et al., 2014). However, both approaches are problematic. If the neural mechanisms that underlie our free decisions are deterministic, then it clearly precludes free will, at least for incompatibilists. By contrast, if the neural underpinnings of free choice are taken to be indeterministic, this provides evidence in favor of free will from a libertarian’s point of view. However, some would argue that indeterminism is just as problematic as determinism because randomness is not the same as free will. The conclusions and interpretation of these experimental data have given rise to several objections.

First, the experimental setup used to investigate free will has its limitations. We will go into these criticisms in detail at the end of the second section. Second, neuroscience may not hold the definitive answer regarding the fundamental nature of brain processes (Bode et al., 2014; Roskies, 2006, 2010, 2022). Indeed, even if some neuronal events appear to be deterministic or indeterministic at a certain level, the complexity of brain wiring and the epistemic limitations of neuroscientific techniques preclude us from drawing any definitive conclusion (Roskies, 2006, 2010, 2012, 2022; Shadlen and Roskies, 2012). Even if determinism is an empirical question, the answer might lie in fundamental physics rather than in neuroscience (Roskies, 2010). Overall, it would seem that neuroscience cannot provide us with a definitive answer regarding the question of determinism and free will. By contrast, a compatibilist’s account of free will requires that we act according to our desires and reasons. Studying decision-making in healthy individuals and in individuals with disorders such as addiction, compulsion, or delusion could provide insights into the brain mechanisms that are required to act according to our desires. This could shed light on the compatibilist kind of free will and could therefore contribute to the field of law and responsibility (Sinnott-Armstrong, 2022). Indeed, legal and moral responsibility can be considered as independent from free will

understood as the ability to do otherwise (Bigenwald and Chambon, 2019; Yaffe, 2022). In a famous example from Frankfurt (1969), a person has a choice between two options A and B. However, a device is installed in their brain, such that if the person chooses A on their own nothing happens, but if the person chooses B then the device is activated and makes the person choose A instead. It is argued that in the situation where the person chooses A on their own they can be held morally responsible for their choice even though they could not have done otherwise (because if the person had chosen B the device would have been activated making them choose A). Thus, what is relevant for moral responsibility is a compatibilist account of free will: that the person acted intentionally, made the choice based on their reasons, was the source of the action, and had sufficient insight to consider the consequences of the action (Bigenwald and Chambon, 2019; Meynen, 2010; Yaffe, 2022). For instance, it is often considered that mental disorders compromise legal responsibility, because the affected patient cannot be considered as the genuine source of their actions (Meynen, 2010). The input from neuroscience could be to provide an assessment of the mental states that are relevant for this definition of legal responsibility (Aharoni et al., 2008): firstly, by evaluating the ability of the person to act intentionally and for a reason; secondly, by assessing whether they can be considered as the source of the action; and, thirdly, by determining the person’s ability to know the nature and quality of the act and to know that the act is wrong, a faculty which, when absent, can lead to the qualification of “insanity”. Using the right experimental task, neuroimaging studies can explore the regions of the brain that are necessary for intentional action, for awareness of intention, and for the formation of moral judgment (Aharoni et al., 2008). An altered activity or structure in any of these regions could be taken as an argument for a defendant to be considered as non-responsible. However, these approaches suffer from important limitations that will be described in further detail at the end of Section 2. Of note, most of the aforementioned studies are based on correlations between brain activity and a particular cognitive function. In consequence, we still lack definitive evidence that the identified brain regions are necessary for the functions required to establish moral responsibility. The current state of neuroscience is thus unable to provide a reliable way of assessing whether a particular individual meets the required conditions for legal responsibility (Aharoni et al., 2008; Bigenwald and Chambon, 2019).

3. Free will and mental causation

3.1. Philosophical questions

We will now discuss the second aspect of free will: mental causation. Mental causation signifies that mental states can cause physical events. Consider that you want to grab a glass of water on the table because you are thirsty. In such a case, we experience the fact that our conscious intention, the phenomenal “feel” associated with this intention, seems to be the cause of our movement toward the glass. This subjective experience or phenomenal feel is of great importance as it is the foundation of our feeling of free will (Fried et al., 2017; Haggard, 2008; Hallett, 2007, 2016). Phenomenal consciousness is thought to have an intrinsic value that adds an important degree of freedom to our behaviors (Cleeremans and Tallon-Baudry, 2022). Conversely, this definition of consciousness has also been criticized (Hacker, 2012). This topic is not new in philosophy of mind, and many authors have pointed out that mental causation is a crucial feature for human agency and free will (Chalmers, 1996; Fodor, 1989; Harbecke, 2008; Kim, 2005; Pockett et al., 2009; Shields, 2014). Because mental and physical properties are of a different nature, the main question that arises is how a mental event can cause a physical event. The philosophical solutions that have been proposed to answer the issue of mental causation strongly depend on one’s position regarding mind-brain ontology, i.e., the nature of the relation between the mind and the physical (Fig. 2) (Harbecke, 2008). We will thus clarify the position we assume regarding this philosophical question. Like

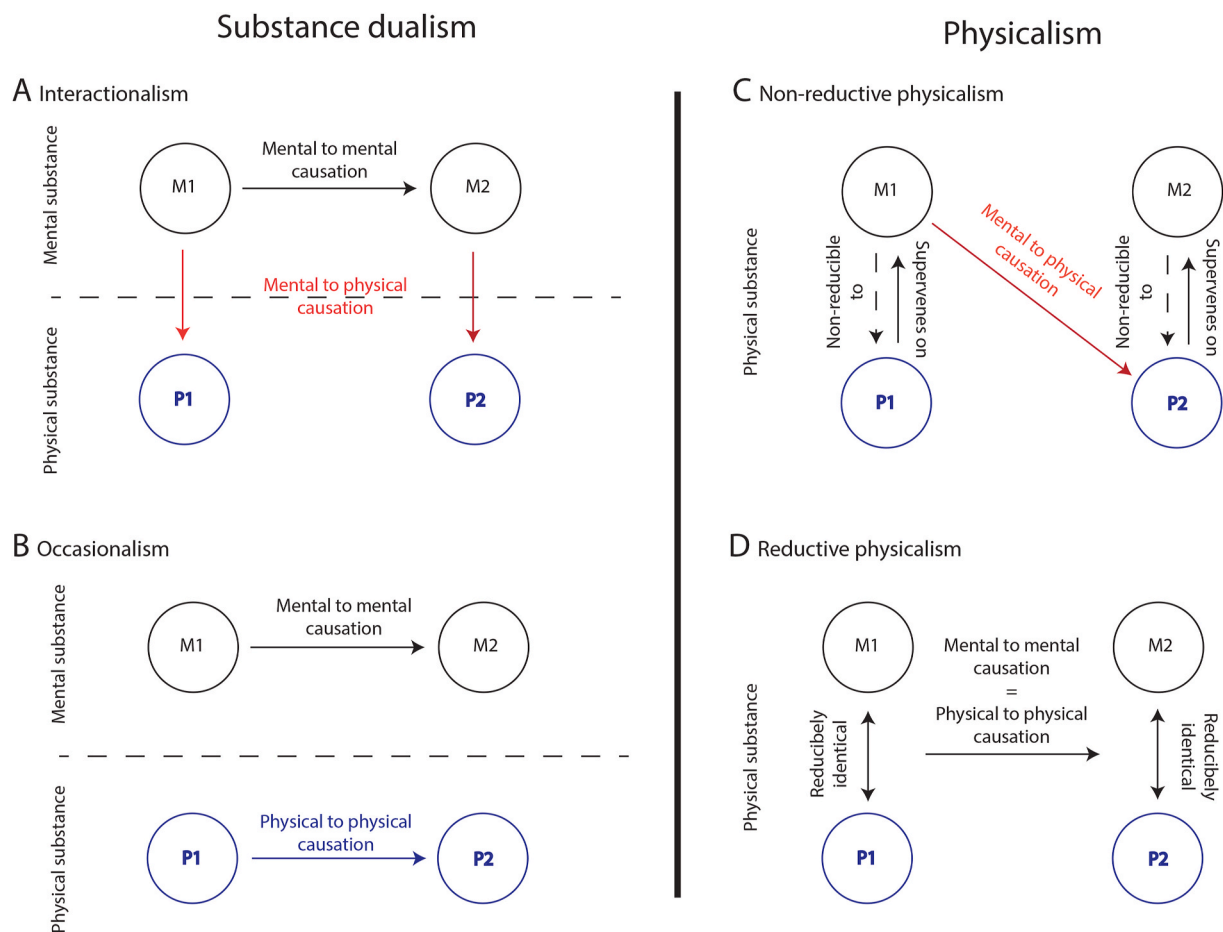


Fig. 2. Different ontological conceptions of the mind-body problem.

Substance dualism refers to an ontology in which the mind and the body are distinct substances with different properties. (A) Interactionism is a form of substance dualism where immaterial minds can interact with physical bodies, and thus cause effects in the physical world. (B) In occasionalism, there is no interaction between the mind and the body, and the coherence between the two is the result of God's action.

In contrast to substance dualism, physicalism states that the universe is composed of bits of matter and their aggregates, and that there is no such thing as immaterial mind, soul or spirit. (C) Non-reductive physicalism is a type of substance physicalism that embraces mind-body supervenience, i.e., the idea that mental events are determined by and dependent on physical events in the brain. However, this position states that mental events are not reducible to physical events. (D) Reductive physicalism, by contrast, considers that mental events are reducible to physical events. Mental events are "nothing more" than physico-chemical events occurring in the brain.

M1: mental event 1; M2: mental event 2; P1: physical event 1; P2: physical event 2.

nearly all neuroscientists and many philosophers, we accept physicalism. This ontology states that the universe is composed uniquely of atoms, molecules, and their aggregates. There is no such thing as immaterial souls or spirits. In addition, many, if not all, physicalists embrace mind-body supervenience, i.e., "the claim that what happens in our mental life is wholly dependent on, and determined by, what happens with our bodily process" (Kim, 1993, 2005). Mental properties are grounded in the physical world, and the relation between the two goes beyond a mere covariation. Mental events such as intentions are dependent on, and determined by, physical events in the brain.

Some have argued that mental causation entails a form of substance dualism, as it seems to suggest that our intentions are driven by non-physical substances (spirits or souls) that are distinct from our body but can nevertheless cause them to move (Haggard, 2008; Haggard and Libet, 2001; Seghezzi and Haggard, 2022). This view was inherited from Descartes's substance dualism, considering minds as immaterial substances that can interact with the physical brain to cause an action (Fig. 2A) (Popper et al., 1977). While being rather intuitive, this concept would preclude mental causation. Indeed, how could an immaterial substance with no spatial extension enter in causal relation with a

material substance (Harbecke, 2008; Kim, 2005)? In Descartes' time, Princess Elisabeth of Bohemia famously argued: "How can the mind of man determine the bodily spirits in producing voluntary actions, being only a thinking substance?" (Garber, 2000). Occasionalism, another form of substance dualism defended by Malebranche, denies any causal relation between the mind and the body. Rather, the two substances exist in parallel, and the apparent causation from the mental to the physical is due to God (Fig. 2B). If substance dualism is ruled out, we are left with two forms of monism, i.e., the view that there is only one kind of substance: absolute idealism and ontological physicalism. Absolute idealism is the idea that the nature of reality is spiritual. Although progress in science has rendered this position less popular, it has its defenders (Hutto, 2000). By contrast, physicalism is the view that the universe is uniquely composed of atoms and their aggregates. Within this physicalist framework, there are two main philosophical positions that account for mental causation. On one hand, according to reductive physicalism, at least some mental properties are reducible to physical properties (Kim, 2005). This means that mental events are "nothing more" than physical events in the brain such as electric currents and the release of synaptic neurotransmitters (Fig. 2D). On the other hand,

non-reductive physicalism (or emergentism) states that the mind supervenes on the physical, but that mental properties are not reducible to physical properties (Chalmers, 1996; J. Fodor, 1997; Putnam, 1975; Searle, 2001). According to this view, although there is nothing more than physical matter, mental events cannot be explained purely by the description of physical events occurring in the brain (Fig. 2C). Reductive and non-reductive physicalism are both committed to ontological physicalism (the fact that there are only physical substances). Still, one of them rejects the idea that mental properties are reducible to physical properties, while the other one embraces it. Both of them consider mental causation as a crucial feature. In the case of reductive physicalism, mental causation is straightforward. Because mental properties are reducible to physical properties, causation from a mental event (the intention to raise my arm) to a physical event (the actual movement of my arm) comes down to causation from a physical event (the brain activity that determines my intention to raise my arm) to another physical event (the actual movement of my arm) (Kim, 2005). By contrast, considering mental causation from the point of view of non-reductive physicalism is more challenging. Indeed, because mental properties emerge from physical properties but are not reducible to these physical properties, mental causation entails a “downward causation” from the higher-level emerging states (such as mental states) to the lower-level physical states (Harbecke, 2008; Kim, 1993; Walter and Heckmann, 2003). However, according to the “exclusion argument” non-reductive physicalism and mental causation are incompatible (Fig. 3) (Kim, 1993, 2005). To reach his conclusion, Kim uses two other principles that both reductive and non-reductive physicalists would willingly accept: (i) the causal closure of the physical: any physical event that has a cause at a moment “t” has a sufficient physical cause at “t”; and (ii) the exclusion principle: a physical event cannot have two distinct and sufficient causes at the same time. Briefly, the exclusion argument goes as follows: let us suppose that a mental event “M” causes a physical event “e” at the instant “t”. Because of the causal closure, “e” also has a sufficient physical cause “P” at “t” (P is the physical supervenience base of M). As a premise of non-reductive physicalism, “M” and “P” are distinct, and thus, “e” has two distinct and sufficient causes at “t”: “M” and “P”. But given the exclusion principle, this is not possible, and hence, “M” has to

be discarded as a sufficient cause for “e”. Indeed, if we discard “P” as a cause of “e”, then “e” would have M as a sufficient cause. Still, because M is a mental event that is irreducible to a physical event, this would violate the causal closure of the physical. This argument leaves us with only two possibilities. On one hand, if one maintains that the mental is irreducible to the physical, mental causation is impossible, and mental events would be purely epiphenomenal. This means that, in the same way as a shadow, mental events are caused by an underlying physical event, but they are themselves deprived of any causal power. On the other hand, embracing reductive physicalism would save mental causation (Kim, 2005), and thus, like many neuroscientists, we favor reductive physicalism (Hopkins et al., 2022). However, this line of reasoning did not resolve the debate, and the validity of the exclusion argument remains a topic of discussion (Walter and Heckmann, 2003). For instance, the formulation of the exclusion principle has been criticized using a different account of causation as “difference-making” (List and Menzies, 2009). Regarding the causal closure of the physical, the current theories in quantum physics argue against the existence of a sufficient physical cause for every physical event (Conway and Kochen, 2006).

3.2. Inputs from neuroscience

Even if we decide to set aside the debate between reductive and non-reductive physicalism and take mental causation for granted, a new “threat” has arisen from neuroscience. Over the last decades, many studies have investigated the phenomenal aspect of volition, meaning the subjective and conscious experience associated with intention. Indeed, conscious intentions are accompanied by a specific subjective experience, a phenomenal feel that subjects can report, and that can thus be studied in different aspects: its relative timing with electrophysiological signals related to movement; its manipulation by altering feedback processing, and its neural substrates. These experiments are thought to give important insights as to whether our conscious intentions can cause our movements. Many neuroscientists claim that the results of the experiments described in the following section demonstrate that conscious intentions are an illusion deprived of any causal

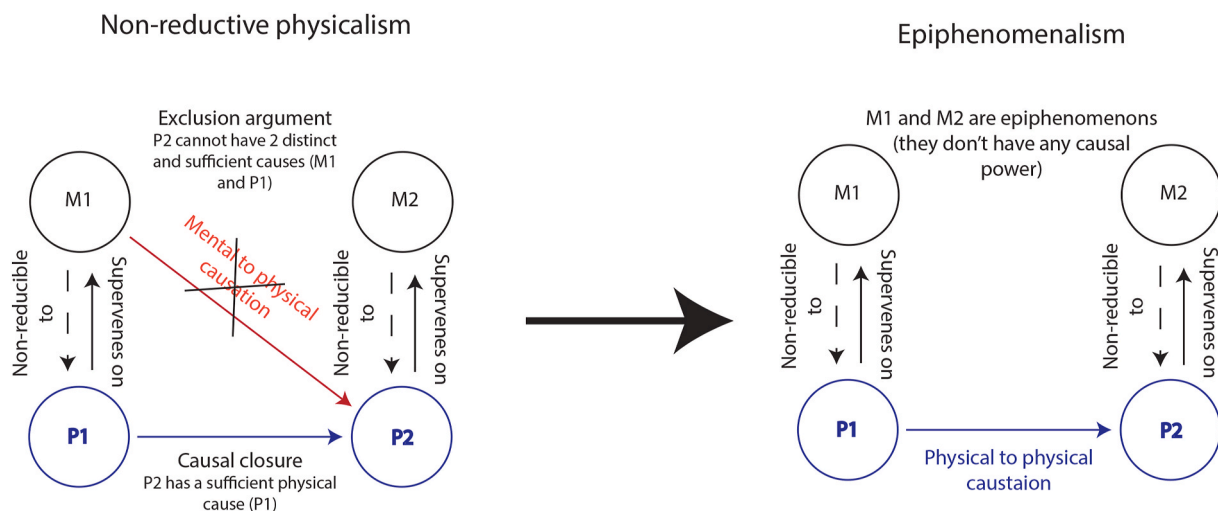


Fig. 3. The exclusion argument.

According to this argument, non-reductive physicalism and mental causation are incompatible. Let us consider that a mental event M1 causes a physical event P2. Because of the causal closure of the physical, P2 has a sufficient physical cause, P1, which is the supervenient base of M1. Because M1 is not reducible to P1, P2 has two distinct and sufficient causes: M1 and P1. In virtue of the exclusion argument this is not possible because otherwise P2 would be over-determined. Thus, M1 has to be rejected as a sufficient cause for P2. According to this argument, non-reductive physicalism precludes mental causation and leads to epiphenomenalism, the fact that mental events are deprived of causal powers.

M1: mental event 1; M2: mental event 2; P1: physical event 1; P2: physical event 2.

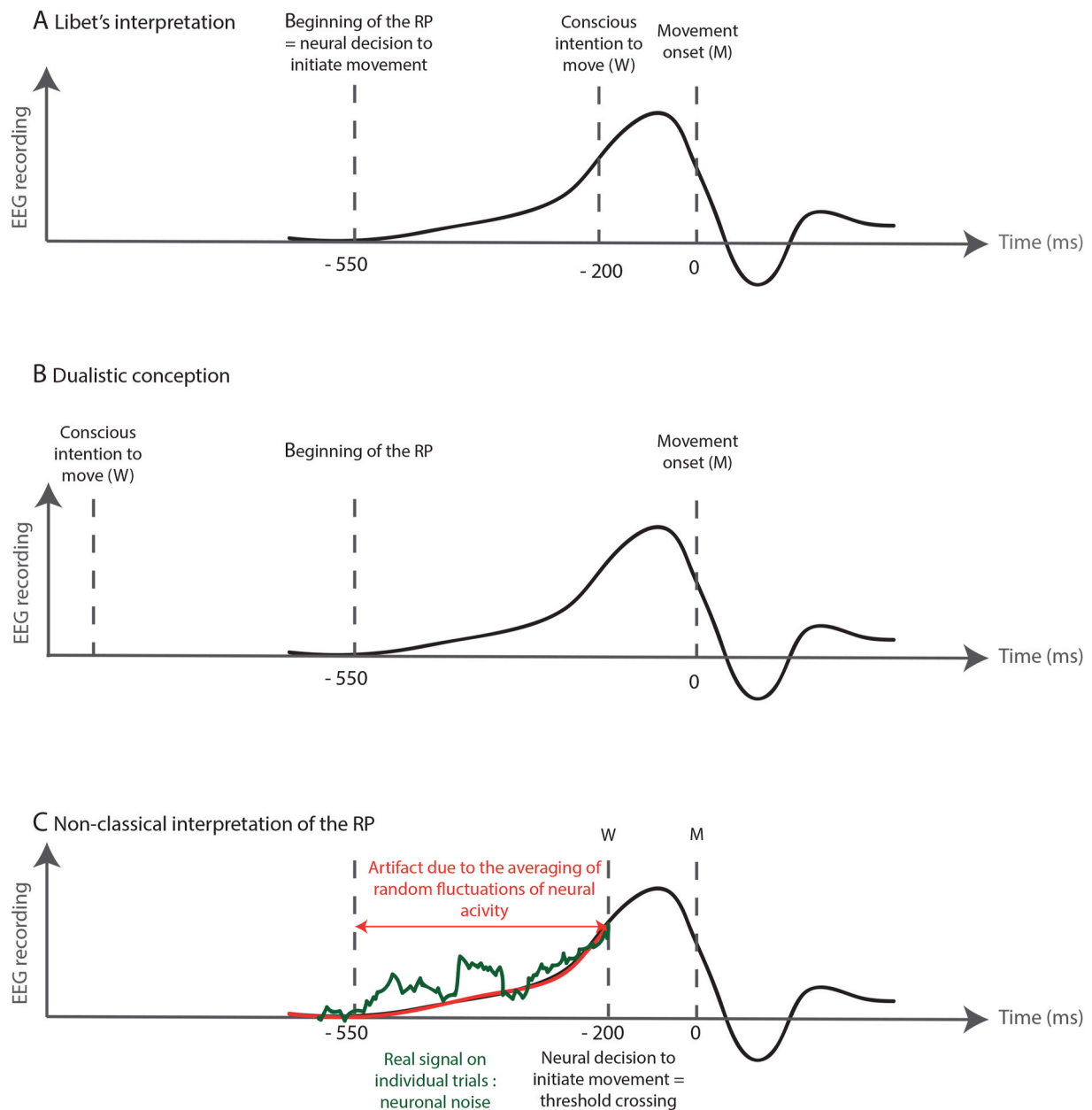


Fig. 4. Different theoretical possible outcomes of Libet's experiments.

(A) According to Libet's view, a physicalist's position implies that the readiness potential (RP) precedes the conscious intention to move, because a mental state can only be considered as a consequence, and not a cause, of brain activity. From this perspective, the conscious intention has no causal power in the generation of the movement, and hence free will cannot exist. (B) In a dualist conception, where immaterial minds interact with our brain to cause our movements, we would expect the conscious intention to precede any brain activity. This theoretical position is incompatible with experimental data. (C) According to another physicalist conception, conscious intentions are mental events that are reducible to the neuronal decision process. Thus, mental events are determined by neuronal events. But, in addition, in virtue of their reducibility to a neuronal event, conscious intentions can also be considered as the cause of movement initiation. In this scenario, we would expect the conscious intention to be simultaneous with the neuronal decision to move. The most recent data on Libet's experiment favor this view and suggest that the early phase of the RP (that precedes the conscious intention) is in fact an artifact (red line) due to the averaging of random neuronal noise (green line). By contrast, the relevant neuronal decision to move (that corresponds to a threshold crossing) seems to coincide with the conscious intention to move (W).

EEG: electroencephalography; RP: readiness potential; W: conscious intention to move; M: movement initiation.

power, or in other words, an epiphenomenon (Wegner and Wheatley, 1999).

3.3. Results from neuroscience

3.3.1. Electrophysiology and timing of conscious intention

The first challenge to the intuitive notion that our conscious intentions are causally efficacious came from Benjamin Libet's

experiments in the 80's (Libet et al., 1983). In this seminal study, participants had to perform a pre-defined movement (a flexion of the wrist or fingers) at a moment they were free to choose. As discussed above, this is a way of operationalizing free will as self-initiated or spontaneous movements, the participants being free to decide the "when" component only. The authors then measured the relative timing of different events. The beginning of the movement, which is called (M), was determined using an EMG recording of corresponding arm muscles. During the

experiment, the participants were looking at a revolving spot on a clock, and they had to memorize the precise location of the spot at the moment they became aware of their intention to move and report it after the movement. This is the instant (W), or the “time of conscious intention to act”. An EEG recorded the brain activity related to the movement. As mentioned in the first section, voluntary movements are preceded by a slow ramping negativity called the RP that is thought to reflect a specific preparatory signal leading up to the movement (Kornhuber and Deecke, 1965; Shibasaki and Hallett, 2006). Libet found that the conscious intention to move (W) preceded movement onset (M) by 200 ms. More surprisingly, the beginning of the RP occurred several hundreds of milliseconds before W (350 ms in Libet’s study, Fig. 4A). These results have been replicated, and a recent meta analysis confirmed that in experiments with a similar setup, the conscious intention to move (W) occurs after the onset of the brain activity (beginning of the RP), with an average delay of around 500 ms (Braun et al., 2021). A similar task conducted in epileptic patients, where the brain activity in the medial frontal areas could be precisely and directly recorded during the task with intracranial electrodes, led to the same conclusion (Fried et al., 2011; Haggard, 2011).

Thus, it would seem that the beginning of the brain activity causing movement (the RP) starts before the participants become conscious of their intention to move. Using the words of Libet: “the brain “decides” to initiate or, at least, to prepare to initiate the act before there is any reportable subjective awareness that such a decision has taken place” (Libet, 1985). This temporal sequence would preclude that the participants’ conscious intention can be the cause of the RP, leading to the conclusion that the conscious intention could not be the cause of the movement (Haggard, 2008; Haggard and Libet, 2001; Hallett, 2007, 2016; Libet et al., 1983; Sinnott-Armstrong and Nadel, 2011). These results have been interpreted as evidence that conscious intention is an epiphenomenon that is deprived of any causal power regarding the initiation of our movements, thus threatening our commonsense conception of free will (Haggard and Libet, 2001; Libet, 1985; Libet et al., 1983; Wegner and Wheatley, 1999).

Another set of experiments using online EEG analysis led to similar conclusions (Bai et al., 2011; Schneider et al., 2013). Using another EEG signal, namely an event-related desynchronization in the beta frequency bands and a learning algorithm, the authors were able to make online predictions about a forthcoming movement more than 0.5 s before the beginning of the muscle activity recorded by EMG (Bai et al., 2011). The main difference between this experiment and Libet’s study is that there was no need to average the EEG data across several trials: the prediction was performed at a single-trial level. Using the same setup, the authors then asked the subjects what they were thinking about at the time of the prediction, before movement onset. While they did sometimes report an intention to move, sometimes they did not (Schneider et al., 2013). This further suggests that movements can be initiated unconsciously.

3.3.2. fMRI and the timing of conscious intention

Libet’s findings have been replicated in a fMRI experiment (Soon et al., 2008). Participants performed a Libet task while their brain activity was measured with fMRI. They were asked to freely choose between right- and left-finger movements when they felt the urge to do so. Participants watched a stream of letters on a screen and had to memorize the letter that was displayed at the moment they became conscious of their intention to move. Using statistical pattern recognition methods, the authors showed that information predicting the forthcoming movement was encoded in frontal and parietal regions as early as 8 s before the subjects reported their conscious intention to move (Soon et al., 2008). These findings were then replicated (Bode et al., 2011) and expanded to abstract decisions where participants were free to choose whether to add or subtract numbers (Soon et al., 2013). In accordance with Libet’s findings, the authors concluded that unconscious brain activity can precede and influence free choices several seconds before they reach consciousness.

3.3.3. Conscious intention and feedback processing

Marc Jeannerod suggested that conscious intentions arise, at least in part, from the processing of the sensory feedback generated by actions (Jeannerod, 2009). The following experiments are based on how the manipulation of feedback from the movement affects subjective reports about the conscious intention to move. In a visuomotor task, healthy participants were asked to draw a straight line on a graphic tablet. The subjects were deprived of direct visual feedback of their own hand, but they could see the line they were drawing on a screen. In some trials, the line was deviated as compared to their actual movement, thus creating a visuomotor conflict (Fournier and Jeannerod, 1998). When the deviation was mild, the participants unconsciously corrected their movement to produce a straight line on the screen, i.e., compensating the erroneous visual feedback. At higher levels of deviation, the participants suddenly became aware of the deviation and consciously intended to correct the trajectory of the line on the graphic tablet (Fournier and Jeannerod, 1998). A haptically deafferented patient who was physiologically deprived from proprioceptive feedback performed the same task. This patient was unable to report any conscious intention to correct her movements, or any conscious intention about her own acts (Fournier et al., 2002). Under such an experimental setting, the conscious intention arises from the comparison between the expected movement outcome and the actual sensory feedback generated by the action (Fournier et al., 2002; Fournier and Jeannerod, 1998; Jeannerod, 2009). The forward model is central to this conception of motor control. According to this model, when performing an action, a copy of the motor command (called the “efference copy”) is used to predict the sensory consequence of the movement. Brain regions (called “comparators”) compare the efference copy with the sensory feedback generated by the movement (Miall et al., 1993; D. Wolpert et al., 1995; D. M. Wolpert and Ghahramani, 2000). In the visuomotor task described above, the participants suddenly became conscious of their corrective movements when they detected a discrepancy between the expected and actual outcome of the movement. These results suggest that conscious intention is a *post hoc* phenomenon that is generated by the processing of the sensory feedback from our action, thus reversing the intuitive causal relation between intention and action (Eagleman, 2004; Jeannerod, 2009).

In keeping with these results, it has been shown that manipulation of the sensory feedback generated by a movement could influence the reported time of conscious intentions. For instance, in a variant of Libet’s study, the participant’s movement was followed by an auditory tone with a variable delay (Banks and Isham, 2009; Rigoni et al., 2010). It was shown that W was delayed linearly with the delay between the movement onset and the tone. In other words, a delay in the perceived time of the action’s outcome resulted in a delay in the reported time of conscious intention. In most of the trials, the subjects even reported W to occur after the actual beginning of the movement. This demonstrates that processes related to movement execution and feedback processing can modulate the reported time of conscious intention.

In a TMS experiment, a pulse applied to the preSMA during or just after action execution could influence the perceived onset of intention. Conscious intention, which supposedly occurs before the movement, may thus be influenced by a brain event that occurs after action execution (Lau et al., 2007). Similarly, an EEG study using tDCS applied to the left angular gyrus or to the primary motor cortex suggested that the reported time of conscious intention is computed from a time period that spans from 1 s before to a few hundreds of milliseconds after movement onset (Douglas et al., 2015).

A recent study investigated the prospective and retrospective factors that contribute to the formation of conscious intentions (Schultze-Kraft et al., 2020). The authors used a brain computer interface and EEG monitoring to detect the RP in real time at a single-trial level. The computer then generated a cue that was triggered by the presence or absence of an RP. The cue could indicate to move or to remain still. The participants were then asked whether they had the intention to move at

the time the cue was displayed. The main finding was that the participants were more likely to report an intention to move when the cue was triggered by an RP and when they had just moved. The authors concluded that conscious intentions rely on both prospective signals linked to motor preparation and on retrospective signals associated with action execution.

All together, these data suggest that the formation of a conscious intention is at least partially retrospective as it relies on a *post hoc* processing of events that occur after action execution. This reverses our intuitive causal relation between intention and action, and is difficult to conciliate with the idea that conscious intentions cause action.

3.3.4. Neuroanatomy of conscious intention

In a physicalist framework, conscious intentions, whether they have a causal power or not, supervene on physical substrates.

In a seminal fMRI study, Lau and colleagues asked the participants to perform a Libet task while focusing on their intentions or on their movements (Lau et al., 2004). Paying attention to a specific cognitive task is thought to increase the hemodynamic signal in the brain regions dedicated to this specific function (Eagleman, 2004), although this hypothesis has been criticized (Nachev and Husain, 2010). Paying attention to intentions resulted in an increased fMRI signal in three areas: the preSMA, the dorsal prefrontal cortex, and the intraparietal sulcus (Lau et al., 2004). These results overlap with data from electrical brain stimulations performed during awake brain surgery. PreSMA stimulation can give rise to an “urge” to move, and increasing stimulation intensity can cause a movement (Fried et al., 1991). Interestingly, movements elicited by preSMA stimulation were identified by the subjects as their own, while the stimulation of the primary motor cortex elicits movements that were perceived as external (Fried et al., 2017). By contrast, stimulation of the inferior parietal lobule resulted in an “intention” to move, but increasing stimulation intensity did not elicit any movement, although some participants reported that they had the feeling they actually moved (Desmurget et al., 2009; Desmurget and Sirigu, 2012). In keeping with these results, anodal tDCS applied to a specific parietal region (the left angular gyrus) to increase its excitability can alter the reported time of movement intention in healthy individuals (Douglas et al., 2015). Furthermore, different disorders have been associated with an altered capacity to report the time of conscious intention. When patients with cerebral lesions in the inferior parietal lobule perform a Libet task, the reported time of their conscious intention (W) is much closer to movement onset as compared with healthy volunteers (Sirigu et al., 2004), as in patients with functional movement disorders (Fried et al., 2017). The latter patients also exhibit a hypoactivation of the same region during abnormal movement (Voon et al., 2010), suggesting a role of the angular gyrus in conscious intentions (Fried et al., 2017).

3.4. Can neuroscience enlighten the philosophical debate regarding mental causation?

3.4.1. Conclusions from neuroscience

These results have led many scientists to conclude that conscious intentions have no causal role regarding the initiation of movement, and that intuition regarding mental causation is in fact an illusion (Haggard, 2008; Hallett, 2007, 2016; Jeannerod, 2009; Libet, 1985; Wegner and Wheatley, 1999). According to this view, conscious intention and action should be seen as two separate, parallel causal chains without any actual causal link between them (Hallett, 2007; Jeannerod, 2009; Wegner and Wheatley, 1999). In keeping with this, conscious intentions have also been conceived as a perception, i.e., a passive phenomenon deprived of causal power regarding the generation of movement (Fried et al., 2017; Hallett, 2007, 2016). The experience of will could be a *post hoc* inference triggered by the movement and relying on a specific set of properties between the intention and the action (Kühn and Brass, 2009; Wegner and Wheatley, 1999). For Wegner, both conscious intentions and the

action itself are produced by unconscious brain mechanisms, and the experience of will is an illusion due to the perception of an apparent causal link between the intention and the action.

From this perspective, what is the alleged role of conscious intention? Libet proposed that although conscious intentions have no causal role in the initiation of action, the delay between W (the time we become conscious of our intentions) and M (the beginning of the movement) was sufficient to allow conscious intention to stop the ongoing preparation of the movement (Libet, 1985). Thus, a conscious intention would exert a “veto power”, and a “free will” would rather be a “free won’t” (Libet, 1985). For others, conscious intention is at the heart of another key feature of human cognition: the sense of agency. Intention, as a conscious representation of the action, includes a representation of the goal to be achieved (Mylopoulos and Pacherie, 2019). The causal role of a conscious intention could thus lie in the comparison between the representation of that goal and the actual consequences of the action. A conscious intention that matches the outcome of a movement results in the subjective experience of being the author of the action, namely, the sense of agency (Haggard, 2017; Hallett, 2007; Jeannerod, 2009).

However, before we accept these radical conclusions, we should carefully examine the criticisms that have been voiced about these studies. Libet’s results, in particular, have given rise to numerous comments. We will review the main arguments that have been put forward to undermine Libet’s conclusions, and we will see that they largely apply to the other studies we previously mentioned.

3.4.2. Criticisms of the conclusions from neuroscience

The various kinds of criticism directed against Libet’s interpretation have been reviewed elsewhere (Brass et al., 2019; Mele, 2009, 2014; Neafsey, 2021; Roskies, 2011; Schurger et al., 2021; Shields, 2014). They can be separated into two main types: technical and philosophical (Roskies, 2011). The technical issues concern the robustness of the data, the method used to report conscious intentions, and the nature of the RP. The philosophical questions tackle the concept of mental causation, the nature of W and the generalizability of the conclusions.

3.4.2.1. Technical comments

3.4.2.1.1. Robustness of the data. A first issue concerns the robustness of the data from which Libet and others have drawn their conclusions. Although Libet’s findings have been replicated, a recent meta-analysis points out that the evidence supporting these results is rather thin. Indeed, the amount of data regarding the difference in timing between unconscious brain activity and conscious intention to move is in fact limited, as it only corresponds to six studies with a total of 53 participants. This limited sample implies that the results should be interpreted with caution (Braun et al., 2021).

3.4.2.1.2. The measure of W. Another question regards the validity of the measure of W as a reflection of the time when the participants become conscious of their intention to move. In Libet’s experiment, the subjects have to remember the position of the spot at the moment they become aware of the decision, but they report it only after they have executed the movement. The cognitive operation of looking at the clock while performing the task and remembering its position might be time-consuming and could lead to measurement bias (Roskies, 2011). As mentioned in the previous section, because W is reported after movement execution, it could be influenced by a brain event that took place after action execution (Banks and Isham, 2009; Douglas et al., 2015; Lau et al., 2007). Another important question is to know whether the timing of W can be influenced by the measurement method itself. Using a variant of the Libet task, Matsushashi and Hallett measured the moment when the subjects started to think of their movement (instant “T”). Rather than using retrospective reports of conscious intention, they introduced a random auditory cue used to probe whether the participants were thinking about the movement at a given time. They showed that the mean time of T was 1.42 s before movement onset, much earlier

than the previous measures of W (Matsushashi and Hallett, 2008). Although on average the beginning of the RP still preceded T, T preceded the RP in four of the 15 subjects, thus undermining Libet's results (Brass et al., 2019; Neafsey, 2021).

3.4.2.1.3. The nature of the readiness potential. One of the main issues regarding Libet's conclusions concerns the interpretation of the RP. An important and implicit assumption made by Libet is that the RP is the result of a decision process that will eventually cause a movement (Brass et al., 2019; Roskies, 2011). In other words, at RP onset the decision to move has already been made unconsciously, and the RP reflects the preparatory process that will generate the movement. This interpretation led Libet to conclude that the conscious intention that follows RP onset has no causal role in initiating the movement (Mele, 2009). However, that RP is a cause of the movement is an assumption that is compatible with the possibility that a conscious intention, even occurring later, is also a cause of the movement. The RP could be a remote cause, while the conscious intention could be a more proximal cause (Mele, 2009). This highlights the fact that Libet implicitly assumes that the RP would have resulted in a movement even if the participants had not acquired a conscious intention to act afterwards. In other words, the RP is supposed to be a *sufficient* cause of the movement.

Direct evidence that the RP is necessary and sufficient to cause a movement is still lacking. An important step would be to show that the RP selectively occurs before a voluntary action, as suggested by a recent work (Travers et al., 2020). However, because the amplitude of the RP is weaker than neuronal noise, this signal is usually obtained from a large number of trials that are time-locked back-averaged to the movement onset. This procedure could lead to important artifacts (Roskies, 2011; Schurger et al., 2021). A more direct method would be to detect RPs on single trials and to show that they can predict subsequent movements. However, both EEG and intracranial recordings failed to provide compelling evidence supporting such a view. The main issue in the study using online EEG recordings to predict movements at a single-trial level (Bai et al., 2011) was accuracy. Sensitivity was below 50%, meaning that the algorithm failed to anticipate more than half of the movements. When a prediction was made, the false positive rate was around 30%, i. e., 30% of the predictions were not followed by an actual movement. Regarding the study using intracranial recordings of single neurons (Fried et al., 2011), although the activity of single neurons averaged over a large number of trials provided a ramping like activity resembling the RP, such a clear signal was not apparent at the single-trial level (Schurger et al., 2021). In addition, RPs can be detected in forms of decision-making that do not require any movement, suggesting that the RP is a more general signal that is not specifically related to movement (Alexander et al., 2016). Conversely, a recent study showed an absence of RP preceding a meaningful decision as compared with an arbitrary decision, although a motor outcome was present in both cases (Maoz et al., 2019). Thus, it would appear that the RP is neither necessary nor sufficient to trigger the action.

An alternative hypothesis could be that the RP represents the process causing the conscious intention, and not the movement. This would seriously undermine Libet's conclusion, because it means that a conscious intention could be the cause of a brain process leading to the movement (Roskies, 2011). However, investigations regarding the association between the RP and conscious intention have resulted in conflicting results. On one hand, the timing of the conscious intention does not correlate with the timing of the RP, making it unlikely that the former is caused by the latter (Haggard and Eimer, 1999; Schlegel et al., 2013). In addition, it was shown that the RP was not influenced by whether the subjects were conscious of the action or not, suggesting once again that the RP is not the cause of the conscious intention (Keller and Heckhausen, 1990; Schlegel et al., 2015). On the other hand, a recent work using a real time EEG analysis method (described in the previous section) showed that subjects were more likely to report an intention to move in the presence of an RP, suggesting a causal link between the two (Schultze-Kraft et al., 2020).

Another hypothesis suggests that the RP could reflect the brain activity related to alternative possible movements that are not executed, a theory known as affordance (Gibson, 1977; Nachev and Hacker, 2014). The brain activity observed when preparing an action might not only be related to the movement that will actually be performed, but also to the alternative possible movements that will not be executed. When choosing "when" to execute a movement (such as in Libet's task), the widening of the distribution of neural activity (the RP) may reflect the alternative possible timing of movement execution. From this perspective, it is impossible to determine at what point the brain activity is causally responsible for the executed movement. Indeed, the neural signal related to the actual movement cannot be disentangled from the signals related to the alternative movements that will not be performed (Nachev and Hacker, 2014).

A final hypothesis could be that the RP might be an artifact due to the experimental instructions. However, a recent EEG study showed that RP also occurred before spontaneous uninstructed movements (Houdayer et al., 2020).

The most serious challenge to this interpretation of the RP as a preparatory signal that is necessary and sufficient to cause voluntary movements came from a study suggesting that it could be an artifact due to the averaging of stochastic neuronal noise (Schurger et al., 2012). In a standard Libet task, the participants performed a pre-defined movement at a time they were free to choose. As described in the first part of this review, the authors hypothesized that the free choice regarding the timing of movement could be described with an accumulation-to-bound model, where neuronal random noise is accumulated as evidence until it reaches a threshold that triggers the decision to move. Using this mathematical model, the authors showed that when random fluctuations at a single-trial level are back-averaged from the time of threshold crossing, it resulted in a signal that had the shape of the early phase of the RP (Schurger et al., 2012). To further test their hypothesis, the authors conducted a second experiment: the participants were interrupted at random times with a cue that prompted them to perform the movement as fast as possible. In accordance with their model, the neuronal random fluctuations were close to the threshold in the trials with the fastest reaction time, while it was far from the threshold in the other cases (Schurger et al., 2012). This view has also been supported from animal experiments (Murakami et al., 2014). The main difference with the classic interpretation is that the RP represents a stochastic process leading to the decision to move, rather than the result of the decision (Schurger et al., 2021). Crucially, the RP onset, classically seen at the beginning of the movement preparation following an unconscious decision to move, could be a mere averaging artifact with no physiological meaning. In addition, this model predicts that the relevant neuronal event regarding the decision to move, namely the threshold crossing, coincides with the timing of conscious intention, as confirmed in another study using a perceptual decision-making task (Kang et al., 2017). In other words, the timing of conscious intention (W) is not the time when we become aware of a decision made unconsciously several hundreds of milliseconds before, but rather the time of the neuronal decision itself (Brass et al., 2019). Overall, this provides a parsimonious explanation that is compatible with our intuition of mental causation. Conscious intentions could be reduced to a neuronal event, the crossing of a threshold that triggers the decision to move, and the preceding early RP may simply reflect an artifact resulting from the averaging of the random fluctuations leading to this decision (Schurger, 2018; Schurger et al., 2021).

3.4.2.2. Philosophical criticisms

3.4.2.2.1. Mental causation and conceptual limits. Although intuitive, the idea that conscious intentions can cause actions raises deep conceptual concerns (Nachev and Hacker, 2014). Among the traditional criteria required to define a causal relation, dependence is of particular importance. If we assume that conscious intentions can cause actions,

we should assume a dependence relation between the two. In particular, causation as “difference-making” entails that intentions make a difference to actions, or in other words, that changing the intention changes the action (List and Menzies, 2009). But conscious intentions are not a necessary condition for intentional actions, for many intentional behaviors are not preceded by conscious intentions. For instance, while driving, most of our movements are “automatic”, in the sense that they are not preceded by a conscious intention. Conversely, conscious intentions are not necessarily followed by actions, because free will lies in the ability to do otherwise. According to this view, free will would precisely entail a lack of causal relation between intention and action (Nachev and Hacker, 2014). Indeed, if actions were causally necessitated by prior intentions, we would lack free will because our actions would be entirely determined by antecedent causes. Free will and mental causation would thus be incompatible. This conceptual concern is thought to seriously undermine the data derived from the experiments addressing this question.

3.4.2.2.2. The nature of *W* and conscious intentions. The first concern is related to the nature of the event that Libet’s experiment measures. *W* is supposed to reflect the time of conscious intention. However, conscious intentions are a mental state (or internal state), meaning that there are no objective or external elements that allow experimenters to determine the precise timing of this event (Lee et al., 2022). Rather, we have to rely on the participant’s subjective report. In addition, the fact that intentions lack phenomenological properties, as proposed by some authors, would make it very difficult to establish their precise timing, and thus to identify the underlying brain events (Nachev and Hacker, 2014).

The second issue is related to the very concept and definition of “conscious intention”. Conscious intention can be understood in many different ways (Chalmers, 1996; Roskies, 2011), which is a source of confusion. It can mean that there is something it is like to form the intention, that the intention has a “phenomenal feel”. However this notion has been criticized, arguing that most experiences, in particular intentions, are deprived of qualitative feel (Hacker, 2012; Nachev and Hacker, 2014). An alternative definition can be borrowed from the field of perceptual experience: “access consciousness”. Access consciousness refers to immediate availability for cognitive process, meaning that perceptual information is broadcast to a global workspace (Block, 2005, 2022). From this prospect, conscious perception are considered as an evidence-accumulation process (Pereira et al., 2022). Different perceptions are competing to access consciousness. This competition is arbitrated by an evidence-accumulation process, where the strongest perceptual representation accesses consciousness once a threshold of evidence is crossed (Block, 2005; Pereira et al., 2022). Whether this concept can be extended to intentions remains an open question. It seems this view would fit partially with the results from Matsuhashi and Hallett (Matsuhashi and Hallett, 2008). To avoid retrospective reporting, a random auditory cue instantaneously probed the subjects’ intentions at different times (instant *T*). Instant *T* occurred much earlier than *W*, a lag that might reflect a period of “probe awareness” (Hallett, 2007; Matsuhashi and Hallett, 2008) when the intention is available for reporting, but has not yet reached access consciousness. However, this conceptual distinction between two kinds of conscious intention is speculative and has not been formally demonstrated (Baars and Laureys, 2005; Cleeremans and Tallon-Baudry, 2022). The nature of consciousness remains elusive, and accordingly there is a large number of diverging definitions, each of them stressing a particular function of consciousness (Block, 2005; Chalmers, 1996). A conceptual work is needed to disentangle these different definitions and to point out which one is relevant for the debate about free will in order to avoid ill-posed problems (Hacker, 2012).

3.4.2.2.3. Generalizability of the results. The third criticism concerns the generalizability of the findings obtained with Libet’s task or other artificial manipulation (Brass et al., 2019; Mele, 2009; Roskies, 2011).

First, there is a conceptual issue concerning the experiments

comparing internally-generated and externally-triggered movements (Nachev, 2010; Nachev and Husain, 2010). It is classically assumed that comparison of the two conditions reveals the neurophysiological substrates of voluntary actions. However, an important and as yet unvalidated assumption is that the two conditions are comparable. Consider the example of a choice between a right-hand and a left-hand movement that is externally-triggered (a visual cue indicates which hand to move) or internally-generated (the cue indicates that the subject can freely choose between the two hands). In the externally triggered condition, the nature and number of factors that influence the outcome is precisely controlled (a left arrow will be associated with a left-hand movement, and vice versa). By contrast, in the free choice condition, the nature and number of internal factors that will lead to the decision is completely unknown. When performing a succession of choices in the absence of any external evidence to determine the answer, the participants may rely on the accumulation of contextual evidence, like the history of past choices, in an attempt to produce a pseudo-randomized behavior (Bode et al., 2014; Nachev, 2010). Importantly, while this kind of information is usually not consciously monitored by the subjects, it can substantially influence the choices made by the participants (Bode et al., 2014). The accuracy of the algorithm proposed by Soon and colleagues to predict the choice of participants as early as 8 s before they acted (Soon et al., 2008) is in fact similar to the accuracy of predictions that are based on the sequence of preceding responses (Lages and Jaworska, 2012). In this study, the subjects selected for the analysis produced pseudo-randomized behaviors (Soon et al., 2008). Rather than decoding the subsequent choice of the subjects, the fMRI patterns detected by this algorithm reflect the bias in the history of past choices that will influence the next one (Bode et al., 2014; Brass et al., 2019; Nachev, 2010). It has thus been argued that the contrast between the two conditions (internally-generated and externally-cued) may not reflect the neuronal networks underlying voluntary actions. Instead, it could be caused by a difference in the number and nature of factors that determine the action (Nachev, 2010; Nachev and Husain, 2010). This conceptual objection cannot be tested with the existing experimental setups, and seriously undermines the results obtained in these studies.

Second, the degree of freedom (the number of possibilities) proposed in these tasks are limited. In Libet’s study, the participants had to choose when to perform a movement, but the movement was pre-determined. In addition, the timing of the movement, supposedly free, had some constraints, for instance regarding the interval between two movements. In the tasks exploring the “what” component of intentional actions, the choice was limited to deciding between right- or left-hand movements.

Third, it is misleading to assume that an internally-generated movement is necessarily more voluntary than an externally-cued action. Some movements can occur in the absence of a clear external trigger, such as yawning or blinking, and it is difficult to say that they reflect free will. Conversely, the first movements of an improvisation dancer that are triggered by the first notes of a piece of music would probably be considered as free.

Fourth, the mere fact that a participant agrees to comply with an experimenter’s demands in the first place, implies that the movements they perform in Libet’s task are not free (Mele, 2009; Pockett et al., 2009; Roskies, 2011). A participant may form conscious intentions about how to perform the task at the beginning of the experiment, which may limit the “freedom” of the individual movements performed during the task. It has been suggested that subjects enter a state of “preparedness” during each trial to introspectively monitor their intentions, and that the detection of an internal signal determined at the beginning of the experiment could trigger the movement. Thus, there would be two kinds of intentions at stake (Brass et al., 2019; Mele, 2009, 2014; Mylopoulos and Pacherie, 2019): a distal conscious intention formed at the beginning of the task, concerning the determination of the internal signal to monitor; and a proximal intention resulting from the detection of such a signal formed before the execution of each movement. In Libet’s tasks, the proximal intentions are the focus of investigation:

whether or not they are indeed unconsciously initiated does not affect our conception of free will, it is rather the formation of distal intentions that is relevant (Brass et al., 2019; Mele, 2009). We cannot generalize from these results that distal intentions are also initiated unconsciously.

Last and most importantly, the nature of the actions that are studied in such experimental setups may not be fully relevant for the free will debate. The proposed movements are both meaningless and arbitrary. In philosophy, these situations are usually referred to as “liberty of indifference”. Rather, we would expect to exert our free will when confronted with a situation where the different alternatives have significant consequences. When we have equally powerful reasons to choose between two sets of an alternative, we are torn between the different possibilities. It is only when facing a “torn decision” that our choices are meaningful, i.e., we are exercising our free will (Kane, 1998). The substantial difference between these two situations can be subsumed under the distinction between “picking” and “choosing” (Brass et al., 2019; Ullmann-Margalit and Morgenbesser, 1977). To take an example from Adina Roskies, when putting on your pants, which leg you put in first is a matter of “picking”, not “choosing”. Whether you put the right or left leg in first makes absolutely no difference, and it would not be surprising that such an action unfolds unconsciously. The argument here is that the movements performed in Libet’s task, and in most of the tasks used in neuroscience, are similar to “picking” which leg to start with when you get dressed. Most of us would accept that such an event unfolds unconsciously, but only a few would take this as evidence that all of our movements and actions are unconsciously initiated in a similar manner, and thus that we lack free will (Roskies, 2011).

To summarize, even if Libet’s results were considered as correct and reproducible, meaning that the movements investigated in this task are initiated unconsciously, it does not necessarily affect our conception of free will. Indeed, we cannot speculate from these findings that other types of actions and decisions, in particular the ones that are relevant for free will, are similarly initiated unconsciously.

At this point, an important objection can be raised. As noted earlier, the neuronal mechanisms underlying complex and arbitrary decisions are probably the same. We described how an accumulation-to-bound model could account for these different kinds of decision-making (Bode et al., 2014; Brass et al., 2019; Murakami et al., 2014; Roskies, 2010; Schurger et al., 2016). Even though the situations studied in laboratories are limited and simplistic, they should still be able to capture the important features of free will (Haggard, 2008). However, as stressed previously, it would seem that a crucial difference between the two situations lies in the nature of the evidence that is accumulated to reach the response threshold. We saw that in the “free-choice” paradigms, the absence of external evidence results in an accumulation of stochastic neuronal noise. By contrast, in real-life situations, our decisions are based on reasons that matter to us. A recent study confirms this idea (Maoz et al., 2019). In this task, the participants had to choose to donate their money to one of two non-profit organizations with two different kinds of decision: arbitrary and deliberate. For the arbitrary decision, the same amount of money was allocated to the two organizations irrespective of the participant’s choice. For the deliberate decision, the participant’s choice determined which of the two organizations received the money. Importantly, in both arbitrary and deliberate decisions, the choice involved a similar movement as the participants had to indicate their response by pressing a button with the left or right hand. Interestingly, an RP was identified for arbitrary decisions, but not for deliberate decisions. The authors concluded that the RP is an artifact resulting from the accumulation of random neuronal noise (Schurger et al., 2012), explaining why it is observed specifically in arbitrary decisions, but not in deliberate decisions (Maoz et al., 2019). In keeping with these results, two other studies suggested that the neural substrates of deliberate and arbitrary decisions differ (Bold et al., 2022; Khalighinejad et al., 2019; Travers et al., 2021). By contrast, other studies failed to find any difference in the RP between arbitrary and deliberate decisions (Bold et al., 2022; Parés-Pujolràs et al., 2021; Travers and

Haggard, 2021; Verbaarschot et al., 2019). Although there is no definitive consensus regarding this issue, the results strongly undermine the generalization of Libet’s results. Indeed, the RP might be a brain signal that is specific to arbitrary decisions, but that would be absent in the real-life situations that are relevant for free will.

3.4.3. Toward a reconciliation of mental causation and experimental data

The way Libet presented his hypothesis played an important role in the apparent incompatibility between neuroscience results and mental causation. Libet and others considered only two possible outcomes of the experiment (Haggard and Libet, 2001) (Fig. 4). On one hand, according to their physicalist view, conscious intentions follow the beginning of the RP, and thus conscious free will cannot exist (Fig. 4A). On the other hand, they argued that the only alternative is substance dualism: mental causation would be possible only if a conscious intention was shown to precede the RP onset (Fig. 4B). In consequence, the authors considered only two possible temporal sequences: either the RP precedes conscious intentions, or conscious intentions precede the RP.

Mental causation and physicalism are thus presented as two incompatible positions. For Libet and others, physicalism implies that mental states are caused by brain activity, while mental causation entails that mental states cause brain activity, a position they present as dualistic. The opposition they draw between mental causation and physicalism comes down to the opposition between dualism and physicalism. However, most contemporary scientists and philosophers would dismiss substance dualism. In addition, different philosophical positions that accept physicalism defend the possibility that mental events are not only the consequence of brain activity, but also the cause of physical events. In particular, reductive physicalism considers that mental events can have physical consequences precisely in virtue of the reducibility of mental events to physical events. So, philosophically, mental causation does not preclude physicalism.

Another possibility that was not considered by Libet is that conscious intentions coincide with the brain signal that causes a movement (Fig. 4C). This assumption would be compatible with physicalism: because mental states supervene on physical states, a mental state cannot precede brain activity, but they can be simultaneous. Crucially, it would also be compatible with the possibility that a conscious intention can cause an action. If conscious intentions are reducible to the neuronal decision to move, in virtue of this reducibility, the causal powers of conscious intentions are identical to those of the underlying brain activity. Thus, conscious intentions could be considered as the cause of movement initiation (Block, 2022; Gavenas et al., 2022). This would reconcile our intuitive notions of free will with the physicalism that many neuroscientists advocate for. We previously saw that recent empirical data and theoretical models support this view by suggesting that W actually corresponds to the moment of the neural decision to move (Brass et al., 2019; Schurger, 2018; Schurger et al., 2021) (Fig. 4C).

Still, the precise neuronal underpinnings of conscious intention need to be investigated (Hopkins et al., 2022), as well as their contribution to action generation (Gavenas et al., 2022; Seth and Bayne, 2022). The respective role of the parietal and prefrontal regions in the generation of conscious intention remains elusive (Desmurget et al., 2009; Desmurget and Sirigu, 2012; Fried et al., 1991). It has been proposed that conscious intention is formed in the posterior parietal cortex, and that the consciously selected response is then planned in the precentral regions (including the SMA and preSMA), before being transmitted to the primary motor cortex (Desmurget and Sirigu, 2012). This view is supported by empirical findings showing an increased connectivity between the parietal and prefrontal regions during voluntary movements (Doganci et al., 2023; Pesaran et al., 2008). Although the precise mechanisms underlying the causal role of conscious intentions in action generation are not completely understood, the theoretical framework we propose allows this possibility.

3.4.4. New experimental paradigms and conceptual limits

The many criticisms we have reviewed necessitate the development of new experimental paradigms to study free will and mental causation. We have already mentioned the study from Maoz et al. which provides a relevant and powerful framework to explore this question (Maoz et al., 2019). Briefly, the participant had to choose to donate money to one of two real non-profit organizations out of a panel of 50. In order to provide an ecological context, real organizations were proposed to the participants, some of which were consensual, and some of which were not. Before the task, the participants were asked to rate the organizations according to their personal preferences. In the deliberate trials, the participant's response resulted in a real donation to the organization they chose, so that their decisions were both meaningful and consequential. Some of the choices were made harder by presenting two organizations for which the participant had a similar preference, in order to reproduce the difficulty of the decisions encountered in real life. Based on this framework, digital immersive environment technologies could provide a computer-based, three-dimensional world, that would allow the experimenter to test complex and realistic decision-making situations (Aharoni et al., 2008).

Still, these situations do not represent the dilemma we face in real life. It is nothing like a "torn decision", where we truly have equally good reasons to choose one option over another (Kane, 1998). For ethical reasons, the consequential decisions that are philosophically relevant for free will cannot be reproduced in a laboratory for the purpose of an experiment. An alternative would be to follow the brain activity of participants in their daily activities with portable devices such as wireless EEG (Barbey et al., 2022; Niso et al., 2023) or portable magnetoencephalography (Boto et al., 2018). With this approach it would be technically possible to investigate mental causation in real-life situations.

These recent advances in technology are promising. However, there are important conceptual limits that go beyond experimental considerations, and which require a true dialogue between neuroscience and philosophy. As outlined before, a conscious intention can be understood in different ways. Philosophers could provide a conceptual clarification regarding which kind of "conscious intention" is relevant when considering mental causation. In the same way, philosophers could enlighten neuroscientists about the different meanings and concepts of causation. What do we mean when we say that intentions cause actions? What kind of dependence relation do we implicitly assume? Philosophers have extensively studied these questions, and neuroscientists should take advantage of this work to produce robust hypotheses that can be properly tested.

4. Conclusion

Free will has been the subject of debate in philosophy for centuries. For a few decades now, neuroscience has also taken a position and claims to bring important insights regarding two elements that are central to the notion of free will: determinism and mental causation. We highlight the advances that have allowed a better understanding of the brain mechanisms underlying decision-making. Although tremendous progress has been made, there is as yet no definitive answer as to whether the brain processes producing our free choices are deterministic or indeterministic. However, reducing the argument to free choice being deterministic or indeterministic presents at least two limitations. First, both possibilities can be considered equally problematic. Determinism is clearly incompatible with at least some forms of free will. However, indeterminism in the form of randomness is also an issue, as it is not synonymous of freedom. Second, because of epistemic limitations, neuroscience might never provide the answer to this question. In the second part, we reviewed the results concerning the causal role of our conscious intentions. The most important challenge came from the observation that conscious intentions seem to follow, rather than to precede, the RP, the brain signal that is thought to cause action.

Following the seminal study of Benjamin Libet, many neuroscientists concluded that conscious intention is deprived of any causal power regarding the initiation of our movements. However, these studies suffer from both technical and philosophical limitations. Importantly, we think it is inappropriate to generalize results obtained with experimental paradigms that may not be fully relevant for investigating free will. In addition, recent data suggest that in Libet's task conscious intention is in fact simultaneous with the neuronal event reflecting the decision, reconciling neuroscience with our intuition that intention can cause action. Thus, there is a need to develop new experimental setups that can measure brain signals related to volition in a more ecological context, reflecting situations with meaningful alternatives and consequences.

Overall, neuroscience may not hold the answer to the longstanding philosophical questions about free will. Nevertheless, neuroscientific input should be considered as important and useful to unravel some of the physical properties of functional brain networks at play when free will is exerted. Neuroscientists and philosophers should foster collaborative and interdisciplinary work and share theoretical and practical experimental knowledge to improve our common understanding of free will. More practically, by further unraveling the brain processes underlying various aspects of free will, neuroscience could enlighten the legal notion of responsibility in the future (Bigenwald and Chambon, 2019; Roskies et al., 2013).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

No data was used for the research described in the article.

References

- Aharoni, E., Funk, C., Sinnott-Armstrong, W., Gazzaniga, M., 2008. Can neurological evidence help courts assess criminal responsibility? Lessons from law and neuroscience. *Ann. N. Y. Acad. Sci.* 1124, 145–160. <https://doi.org/10.1196/annals.1440.007>.
- Alexander, P., Schlegel, A., Sinnott-Armstrong, W., Roskies, A.L., Wheatley, T., Tse, P.U., 2016. Readiness potentials driven by non-motoric processes. *Conscious. Cognit.* 39, 38–47. <https://doi.org/10.1016/j.concog.2015.11.011>.
- Ariani, G., Wurm, M.F., Lingnau, A., 2015. Decoding internally and externally driven movement plans. *J. Neurosci.* 35 (42), 14160–14171. <https://doi.org/10.1523/JNEUROSCI.0596-15.2015>.
- Assal, F., Schwartz, S., Vuilleumier, P., 2007. Moving with or without will : functional neural correlates of alien hand syndrome. *Ann. Neurol.* 62 (3), 301–306. <https://doi.org/10.1002/ana.21173>.
- Baars, B.J., Laureys, S., 2005. One, not two, neural correlates of consciousness. *Trends Cognit. Sci.* 9 (6), 269. <https://doi.org/10.1016/j.tics.2005.04.008> author reply 270.
- Bai, O., Rathi, V., Lin, P., Huang, D., Battapady, H., Fei, D.-Y., Schneider, L., Houdayer, E., Chen, X., Hallett, M., 2011. Prediction of human voluntary movement before it occurs. *Clin. Neurophysiol.* 122 (2), 364–372. <https://doi.org/10.1016/j.clinph.2010.07.010>.
- Balaguer, M., 2009. *Free Will as an Open Scientific Problem*. MIT Press.
- Balaguer, M., 2014. *Free Will*. MIT Press.
- Banks, W.P., Isham, E.A., 2009. We infer rather than perceive the moment we decided to act. *Psychol. Sci.* 20 (1), 17–21. <https://doi.org/10.1111/j.1467-9280.2008.02254.x>.
- Bansal, S., Ford, J.M., Spering, M., 2018. The function and failure of sensory predictions : sensory prediction function and failure. *Ann. N. Y. Acad. Sci.* 1426 (1), 199–220. <https://doi.org/10.1111/nyas.13686>.
- Barbey, F.M., Farina, F.R., Buick, A.R., Danyeli, L., Dyer, J.F., Islam, Md N., Krylova, M., Murphy, B., Nolan, H., Rueda-Delgado, L.M., Walter, M., Whelan, R., 2022. Neuroscience from the comfort of your home : repeated, self-administered wireless dry EEG measures brain function with high fidelity. *Front. Digit. Health* 4, 944753. <https://doi.org/10.3389/fgth.2022.944753>.
- Bigenwald, A., Chambon, V., 2019. Criminal responsibility and neuroscience : No revolution yet. *Front. Psychol.* 10, 1406. <https://doi.org/10.3389/fpsyg.2019.01406>.
- Blakemore, S.-J., Oakley, D.A., Frith, C.D., 2003. Delusions of alien control in the normal brain. *Neuropsychologia* 41 (8), 1058–1067. [https://doi.org/10.1016/s0028-3932\(02\)00313-5](https://doi.org/10.1016/s0028-3932(02)00313-5).
- Block, N., 2005. Two neural correlates of consciousness. *Trends Cognit. Sci.* 9 (2), 46–52. <https://doi.org/10.1016/j.tics.2004.12.006>.

- Block, N., 2022. Do conscious decisions cause physical actions? In: Block, N. (Ed.), *Free Will*. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0012>, pp. 95–108.
- Bode, S., He, A.H., Soon, C.S., Trampel, R., Turner, R., Haynes, J.-D., 2011. Tracking the unconscious generation of free decisions using Ultra-high field fMRI. *PLoS One* 6 (6), e21612. <https://doi.org/10.1371/journal.pone.0021612>.
- Bode, S., Murawski, C., Soon, C.S., Bode, P., Stahl, J., Smith, P.L., 2014. Demystifying “free will”: the role of contextual information and evidence accumulation for predictive brain activity. *Neurosci. Biobehav. Rev.* 47, 636–645. <https://doi.org/10.1016/j.neubiorev.2014.10.017>.
- Bok, H., 1998. *Freedom and Responsibility*. Princeton University Press.
- Bold, J. Lyn, Mudrik, L., Maoz, U., 2022. How are arbitrary and deliberate decisions similar and different? In: lyn Bold, J., Mudrik, L., Maoz, U. (Eds.), *Free Will*. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0019> (pp. 165–174).
- Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S.S., Muñoz, L.D., Mullinger, K.J., Tierney, T.M., Bestmann, S., Barnes, G.R., Bowtell, R., Brookes, M.J., 2018. Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* 555 (7698), 657–661. <https://doi.org/10.1038/nature26147>.
- Brass, M., Furstenberg, A., Mele, A.R., 2019. Why neuroscience does not disprove free will. *Neurosci. Biobehav. Rev.* 102, 251–263. <https://doi.org/10.1016/j.neubiorev.2019.04.024>.
- Brass, M., Haggard, P., 2008. The what, when, whether model of intentional action. *Neuroscientist* 14 (4), 319–325. <https://doi.org/10.1177/1073858408317417>.
- Braun, M.N., Wessler, J., Frieze, M., 2021. A meta-analysis of Libet-style experiments. *Neurosci. Biobehav. Rev.* 128, 182–198. <https://doi.org/10.1016/j.neubiorev.2021.06.018>.
- Chalmers, D.J., 1996. *The Conscious Mind: in Search of a Fundamental Theory*. Oxford University Press.
- Cleeremans, A., Tallon-Baudry, C., 2022. Consciousness matters: phenomenal experience has functional value. *Neurosci. Conscious.* (1) <https://doi.org/10.1093/nc/niac007>.
- Conway, J., Kochen, S., 2006. The free will theorem. *Found. Phys.* 36 (10), 1441–1473. <https://doi.org/10.1007/s10701-s10701-006-9068-6>.
- Cui, H., Andersen, R.A., 2007. Posterior parietal cortex encodes autonomously selected motor plans. *Neuron* 56 (3), 552–559. <https://doi.org/10.1016/j.neuron.2007.09.031>.
- Deiber, M.-P., Passingham, R.E., Colebatch, J.G., Friston, K.J., Nixon, P.D., Frackowiak, R.S.J., 1991. Cortical areas and the selection of movement: a study with positron emission tomography. *Exp. Brain Res.* 84 (2), 393–402. <https://doi.org/10.1007/BF00231461>.
- Dennett, D.C., 2015. *Elbow room*. In: *The Varieties of Free Will Worth Wanting*, new edition. MIT Press.
- Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C., Sirigu, A., 2009. Movement intention after parietal cortex stimulation in humans. *Science* 324 (5928), 811–813. <https://doi.org/10.1126/science.1169896>.
- Desmurget, M., Sirigu, A., 2012. Conscious motor intention emerges in the inferior parietal lobule. *Curr. Opin. Neurobiol.* 22 (6), 1004–1011. <https://doi.org/10.1016/j.conb.2012.06.006>.
- Dirnberger, G., Fickel, U., Lindinger, G., Lang, W., Jahanshahi, M., 1998. The mode of movement selection. Movement-related cortical potentials prior to freely selected and repetitive movements. *Exp. Brain Res.* 120 (2), 263–272. <https://doi.org/10.1007/s002210050400>.
- Doganci, N., Iannotti, G.R., Ptak, R., 2023. Task-based functional connectivity identifies two segregated networks underlying intentional action. *Neuroimage* 268, 119866. <https://doi.org/10.1016/j.neuroimage.2023.119866>.
- Douglas, Z.H., Maniscalco, B., Hallett, M., Wassermann, E.M., He, B.J., 2015. Modulating conscious movement intention by noninvasive brain stimulation and the underlying neural mechanisms. *J. Neurosci.: The Official Journal of the Society for Neuroscience* 35 (18), 7239–7255. <https://doi.org/10.1523/JNEUROSCI.4894-14.2015>.
- Eagleman, D.M., 2004. Neuroscience: the where and when of intention. *Science* 303 (5661), 1144–1146. <https://doi.org/10.1126/science.1095331>.
- Ebert, J.P., Wegner, D.M., 2011. Mistaking randomness for free will. *Conscious. Cognit.* 20 (3), 965–971. <https://doi.org/10.1016/j.concog.2010.12.012>.
- Filevich, E., Vanneste, P., Brass, M., Fias, W., Haggard, P., Kühn, S., 2013. Brain correlates of subjective freedom of choice. *Conscious. Cognit.* 22 (4), 1271–1284. <https://doi.org/10.1016/j.concog.2013.08.011>.
- Fodor, J.A., 1997. *Special sciences: still autonomous after all these years*. *Phil. Perspect.* 11, 149–163.
- Fodor, J.A., 1989. Making mind matter more. *Phil. Top.* 17 (1), 59–79.
- Forstmann, B.U., Brass, M., Koch, I., von Cramon, D.Y., 2006. Voluntary selection of task sets revealed by functional magnetic resonance imaging. *J. Cognit. Neurosci.* 18 (3), 388–398. <https://doi.org/10.1162/0899892906775990589>.
- Fourneret, P., Jeannerod, M., 1998. Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia* 36 (11), 1133–1140. [https://doi.org/10.1016/S0028-3932\(98\)00006-2](https://doi.org/10.1016/S0028-3932(98)00006-2).
- Fourneret, P., Paillard, J., Lamarre, Y., Cole, J., Jeannerod, M., 2002. Lack of conscious recognition of one’s own actions in a haptically deafferented patient. *Neuroreport* 13 (4), 541–547.
- Frankfurt, H.G., 1969. Alternate possibilities and moral responsibility. *J. Philos.* 66 (23), 829–839.
- Frankfurt, H.G., 1971. Freedom of the will and the concept of a person. *J. Philos.* 68 (1), 5–20. <https://doi.org/10.2307/2024717>.
- Fried, I., Haggard, P., He, B.J., Schurger, A., 2017. Volition and action in the human brain: processes, pathologies, and reasons. *J. Neurosci.* 37 (45), 10842–10847. <https://doi.org/10.1523/JNEUROSCI.2584-17.2017>.
- Fried, I., Katz, A., McCarthy, G., Sass, K.J., Williamson, P., Spencer, S.S., Spencer, D.D., 1991. Functional organization of human supplementary motor cortex studied by electrical stimulation. *J. Neurosci.: Offic. J. Soc. Neurosci.* 11 (11), 3656–3666. <https://doi.org/10.1523/JNEUROSCI.11-11-03656.1991>.
- Fried, I., Mukamel, R., Kreiman, G., 2011. Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* 69 (3), 548–562. <https://doi.org/10.1016/j.neuron.2010.11.045>.
- Garber, D., 2000. *Descartes Embodied: Reading Cartesian Philosophy through Cartesian Science*. Cambridge University Press.
- Gavenas, J., Hallett, M., Maoz, U., 2022. Which neural mechanisms could enable conscious control of action? In: Gavenas, J., Hallett, M., Maoz, U. (Eds.), *Free Will*. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0025>, pp. 222–231.
- Gibson, J.J., 1977. In: Shaw, R., Bransford, J. (Eds.), *Perceiving, acting, and knowing*, 67. Lawrence Erlbaum, Hillsdale, pp. 67–82.
- Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574. <https://doi.org/10.1146/annurev.neuro.29.051605.113038>.
- Hacker, P.M.S., 2012. The Sad and Sorry History of Consciousness: Being, Among Other Things, a Challenge to the ‘Consciousness-Studies Community’, vol. 70. Royal Institute of Philosophy Supplements, pp. 149–168. <https://doi.org/10.1017/S1358246112000082>.
- Haggard, P., 2008. Human volition: towards a neuroscience of will. *Nat. Rev. Neurosci.* 9 (12), 934–946. <https://doi.org/10.1038/nrn2497>.
- Haggard, P., 2011. Decision time for free will. *Neuron* 69 (3), 404–406. <https://doi.org/10.1016/j.neuron.2011.01.028>.
- Haggard, P., 2017. Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18 (4), 196–207. <https://doi.org/10.1038/nrn.2017.14>.
- Haggard, P., 2019. The neurocognitive bases of human volition. *Annu. Rev. Psychol.* 70 (1), 9–28. <https://doi.org/10.1146/annurev-psych-010418-103348>.
- Haggard, P., Eimer, M., 1999. On the relation between brain potentials and the awareness of voluntary movements. *Exp. Brain Res.* 126 (1), 128–133. <https://doi.org/10.1007/s002210050722>.
- Haggard, P., Libet, B., 2001. Conscious intention and brain activity. *J. Conscious. Stud.* 8 (11), 47–64.
- Hallett, M., 2007. Volitional control of movement: the physiology of free will. *Clin. Neurophysiol.* 118 (6), 1179–1192. <https://doi.org/10.1016/j.clinph.2007.03.019>.
- Hallett, M., 2016. Physiology of free will. *Ann. Neurol.* 80 (1), 5–12. <https://doi.org/10.1002/ana.24657>.
- Harbecke, J., 2008. *Mental Causation: Investigating the Mind’s Powers in a Natural World*. De Gruyter, Inc.
- Hassan, A., Josephs, K.A., 2016. Alien hand syndrome. *Curr. Neurol. Neurosci. Rep.* 16 (8), 73. <https://doi.org/10.1007/s11910-016-0676-z>.
- Hobbes, T., 1962. *The English Works of Thomas Hobbes, W. Molesworth, London: Scientia Aalen, orig. p. 1654*.
- Hoffstaedter, F., Grefkes, C., Zilles, K., Eickhoff, S.B., 2013. The “what” and “when” of self-initiated movements. *Cerebr. Cortex* 23 (3), 520–530. <https://doi.org/10.1093/cercor/bhr391>.
- Honderich, T., 1990. *A Theory of Determinism*. Clarendon Press.
- Hopkins, A., Mudrik, L., Maoz, U., 2022. How does the absence of a consensus about the neural basis of consciousness and volition affect theorizing about conscious volition? In: Hopkins, A., Mudrik, L., Maoz, U. (Eds.), *Free Will*. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0026> (pp. 232–240).
- Houdayer, E., Lee, S.-J., Hallett, M., 2020. Cerebral preparation of spontaneous movements: an EEG study. *Clin. Neurophysiol.* 131 (11), 2561–2565. <https://doi.org/10.1016/j.clinph.2020.08.002>.
- Hume, D., 1960. *A Treatise of Human Nature*. Clarendon Press.
- Hutto, D.D., 2000. *Beyond Physicalism*. John Benjamins Publishing Company. Aicr.21.
- Hwang, E.J., Dahlen, J.E., Mukundan, M., Komiyama, T., 2017. History-based action selection bias in posterior parietal cortex. *Nat. Commun.* 8 (1) <https://doi.org/10.1038/s41467-017-01356-z>.
- Jeannerod, M., 2009. *Le Cerveau Volontaire*. Odile Jacob.
- Kane, R., 1998. *The Significance of Free Will*. Oxford University Press.
- Kang, Y.H.R., Petzschner, F.H., Wolpert, D.M., Shadlen, M.N., 2017. Piercing of consciousness as a threshold-crossing operation. *Curr. Biol.* 27 (15), 2285–2295.e6. <https://doi.org/10.1016/j.cub.2017.06.047>.
- Keller, I., Heckhausen, H., 1990. Readiness potentials preceding spontaneous motor acts: voluntary vs. involuntary control. *Electroencephalogr. Clin. Neurophysiol.* 76 (4), 351–361. [https://doi.org/10.1016/0013-4694\(90\)90036-J](https://doi.org/10.1016/0013-4694(90)90036-J).
- Khalighinejad, N., Brann, E., Dorgham, A., Haggard, P., 2019. Dissociating cognitive and motoric precursors of human self-initiated action. *J. Cognit. Neurosci.* 31 (5), 754–767. https://doi.org/10.1162/jocn_a_01380.
- Kim, J., 1993. *Supervenience and Mind: Selected Philosophical Essays*. Cambridge University Press.
- Kim, J., 2005. *Physicalism, or Something Near Enough*. Princeton University Press.
- Kornhuber, H.H., Deecke, L., 1965. Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Arch. für Gesamte Physiol. Menschen Tiere* 284 (1), 1–17. <https://doi.org/10.1007/BF00412364>.
- Krainik, A., Duffau, H., Capelle, L., Cornu, P., Boch, A.-L., Mangin, J.-F., Le Bihan, D., Marsault, C., Chiras, J., Lehericy, S., 2004. Role of the healthy hemisphere in recovery after resection of the supplementary motor area. *Neurol. Now.* 62 (8), 1323–1332.

- Krieghoff, V., 2009. Dissociating what and when of intentional actions. *Front. Hum. Neurosci.* 3 <https://doi.org/10.3389/neuro.09.003.2009>.
- Kühn, S., Brass, M., 2009. Retrospective construction of the judgement of free choice. *Conscious. Cognit.* 18 (1), 12–21. <https://doi.org/10.1016/j.concog.2008.09.007>.
- Lages, M., Jaworska, K., 2012. How predictable are « spontaneous decisions » and « hidden intentions »? Comparing classification results based on previous responses with multivariate pattern analysis of fMRI BOLD signals. *Front. Psychol.* 3, 56. <https://doi.org/10.3389/fpsyg.2012.00056>.
- Laplante, D., Talairach, J., Meininger, V., Bancaud, J., Orgogozo, J.M., 1977. Clinical consequences of corticectomies involving the supplementary motor area in man. *J. Neurol. Sci.* 34 (3), 301–314.
- Lau, H.C., Rogers, R.D., Haggard, P., Passingham, R.E., 2004. Attention to intention. *Science* 303 (5661), 1208–1210. <https://doi.org/10.1126/science.1090973>.
- Lau, H.C., Rogers, R.D., Passingham, R.E., 2007. Manipulating the experienced onset of intention after action execution. *J. Cognit. Neurosci.* 19 (1), 81–90. <https://doi.org/10.1162/jocn.2007.19.1.81>.
- Lee, S.J., Wong, S.M., Alice, Maoz, U., Hallett, M., 2022. How can we determine the precise timing of mental events related to action? In: Lee, S.J., Alice, S.M., Wong, U. Maoz, Hallett, M. (Eds.), *Free Will*. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0028> (pp. 251–257).
- Libet, B., 1985. *Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action*, vol. 38. The behavioral and brain sciences 8, 529–566.
- Libet, B., Gleason, C.A., Wright, E.W., Pearl, D.K., 1983. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain* 106 (3), 623–642. <https://doi.org/10.1093/brain/106.3.623>.
- List, C., Menzies, P., 2009. Nonreductive physicalism and the limits of the exclusion principle. *J. Philos.* 106 (9), 475–502.
- Maoz, U., Yaffe, G., Koch, C., Mudrik, L., 2019. Neural precursors of decisions that matter—an ERP study of deliberate and arbitrary choice. *Elife* 8, e39787. <https://doi.org/10.7554/eLife.39787>.
- Matsushashi, M., Hallett, M., 2008. The timing of the conscious intention to move. *Eur. J. Neurosci.* 28 (11), 2344–2351. <https://doi.org/10.1111/j.1460-9568.2008.06525.x>.
- Mele, A.R., 2009. *Effective Intentions : the Power of Conscious Will*. Oxford University Press.
- Mele, A.R., 2014. *Free : Why Science Hasn't Disproved Free Will*. Oxford University Press.
- Meynen, G., 2010. Free will and mental disorder : exploring the relationship. *Theor. Med. Bioeth.* 31 (6), 429–443. <https://doi.org/10.1007/s11017-010-9158-5>.
- Miall, R.C., Weir, D.J., Wolpert, D.M., Stein, J.F., 1993. Is the cerebellum a smith predictor? *J. Mot. Behav.* 25 (3), 203–216. <https://doi.org/10.1080/00222895.1993.9942050>.
- Mochizuki, K., Funahashi, S., 2014. Opposing history effect of preceding decision and action in the free choice of saccade direction. *J. Neurophysiol.* 112 (4), 923–932. <https://doi.org/10.1152/jn.00846.2013>.
- Mueller, V.A., Brass, M., Waszak, F., Prinz, W., 2007. The role of the preSMA and the rostral cingulate zone in internally selected actions. *Neuroimage* 37 (4), 1354–1361. <https://doi.org/10.1016/j.neuroimage.2007.06.018>.
- Murakami, M., Vicente, M.L., Costa, G.M., Mainen, Z.F., 2014. Neural antecedents of self-initiated actions in secondary motor cortex. *Nat. Neurosci.* 17 (11), 1574–1582. <https://doi.org/10.1038/nn.3826>.
- Mylopoulos, M., Pacherie, E., 2019. Intentions : the dynamic hierarchical model revisited. *WIREs Cognitive Science* 10 (2). <https://doi.org/10.1002/wcs.1481>.
- Nachev, P., 2010. *Free choice & voluntary action*. *Psyche Interdiscipl. J. Res. Conscious.* 16 (1), 87–98.
- Nachev, P., Hacker, P., 2014. The neural antecedents to voluntary action : a conceptual analysis. *Cognit. Neurosci.* 5 (3–4), 193–208. <https://doi.org/10.1080/17588928.2014.934215>.
- Nachev, P., Husain, M., 2010. Action and the fallacy of the 'internal' : comment on Passingham et al. *Trends Cognit. Sci.* 14 (5), 192–193. <https://doi.org/10.1016/j.tics.2010.03.002> ; author reply 193–194.
- Nachev, P., Kennard, C., Husain, M., 2008. Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* 9 (11), 856–869. <https://doi.org/10.1038/nrn2478>.
- Neafsey, E.J., 2021. Conscious intention and human action : review of the rise and fall of the readiness potential and Libet's clock. *Conscious. Cognit.* 94, 103171 <https://doi.org/10.1016/j.concog.2021.103171>.
- Niso, G., Romero, E., Moreau, J.T., Araujo, A., Krol, L.R., 2023. Wireless EEG : a survey of systems and studies. *Neuroimage* 269, 119774. <https://doi.org/10.1016/j.neuroimage.2022.119774>.
- O'Connor, D.J., 1972. *Free Will*. Springer.
- O'Connor, T., 2022. What is free will? In: O'Connor, T. (Ed.), *Free Will*. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0005>, pp. 41–48).
- Parés-Pujolràs, E., Travers, E., Ahmetoglu, Y., Haggard, P., 2021. Evidence accumulation under uncertainty—a neural marker of emerging choice and urgency. *Neuroimage* 232, 117863. <https://doi.org/10.1016/j.neuroimage.2021.117863>.
- Passingham, R.E., 1987. Two cortical systems for directing movement. *Ciba Foundation Symposium* 132, 151–164. <https://doi.org/10.1002/9780470513545.ch10>.
- Passingham, R.E., Bengtsson, S.L., Lau, H.C., 2010. Medial frontal cortex : from self-generated action to reflection on one's own performance. *Trends Cognit. Sci.* 14 (1), 16–21. <https://doi.org/10.1016/j.tics.2009.11.001>.
- Pereboom, D., Sosa, E., Dancy, J., Haldane, J., Harman, G., Jackson, F., Lycan, W.G., Shoemaker, S., Thomson, J.J., 2001. *Living without Free Will*. Cambridge University Press.
- Pereira, M., Perrin, D., Faivre, N., 2022. A leaky evidence accumulation process for perceptual experience. *Trends Cognit. Sci.* 26 (6), 451–461. <https://doi.org/10.1016/j.tics.2022.03.003>.
- Pesaran, B., Nelson, M.J., Andersen, R.A., 2008. Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453 (7193), 406–409. <https://doi.org/10.1038/nature06849>.
- Pockett, S., Banks, W.P., Gallagher, S., 2009. *Does Consciousness Cause Behavior?* MIT Press.
- Popper, K.R., Eccles, J.C., Eccles, J.C., 1977. *The Self and its Brain*. Springer.
- Potgieser, A.R.E., de Jong, B.M., Wagemakers, M., Hoving, E.W., Groen, R.J.M., 2014. Insights from the supplementary motor area syndrome in balancing movement initiation and inhibition. *Front. Hum. Neurosci.* 8, 960. <https://doi.org/10.3389/fnhum.2014.00960>.
- The nature of mental states. In: Putnam, H. (Ed.), 1975. *Philosophical Papers : Volume 2 : Mind, Language and Reality*. Cambridge University Press, pp. 429–440. <https://doi.org/10.1017/CBO9780511625251.023>, 2.
- Rigoni, D., Brass, M., Sartori, G., 2010. Post-action determinants of the reported time of conscious intentions. *Front. Hum. Neurosci.* 4, 38. <https://doi.org/10.3389/fnhum.2010.00038>.
- Roskies, A.L., 2006. Neuroscientific challenges to free will and responsibility. *Trends Cognit. Sci.* 10 (9), 419–423. <https://doi.org/10.1016/j.tics.2006.07.011>.
- Roskies, A.L., 2010. How does neuroscience affect our conception of volition? *Annu. Rev. Neurosci.* 33 (1), 109–130. <https://doi.org/10.1146/annurev-neuro-060909-153151>.
- Roskies, A.L., 2011. Why Libet's studies don't pose a threat to free will. In: Nadel, L., Sinnott-Armstrong, W. (Eds.), *Conscious Will and Responsibility. A Tribute to Benjamin Libet*. Oxford University Press, pp. 11–22.
- Roskies, A.L., 2012. How does the neuroscience of decision making bear on our understanding of moral responsibility and free will? *Curr. Opin. Neurobiol.* 22 (6), 1022–1026. <https://doi.org/10.1016/j.conb.2012.05.009>.
- Roskies, A.L., 2022. What kind of neuroscientific evidence, if any, could determine whether anyone has free will? In: Roskies, A.L. (Ed.), *Free Will*. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0009>, pp. 71–79).
- Roskies, A.L., Schweitzer, N.J., Saks, M.J., 2013. Neuroimages in court : less biasing than feared. *Trends Cognit. Sci.* 17 (3), 99–101. <https://doi.org/10.1016/j.tics.2013.01.008>.
- Schlegel, A., Alexander, P., Sinnott-Armstrong, W., Roskies, A., Tse, P.U., Wheatley, T., 2013. Barking up the wrong free : readiness potentials reflect processes independent of conscious will. *Exp. Brain Res.* 229 (3), 329–335. <https://doi.org/10.1007/s00221-013-3479-3>.
- Schlegel, A., Alexander, P., Sinnott-Armstrong, W., Roskies, A., Tse, P.U., Wheatley, T., 2015. Hypnotizing Libet : readiness potentials with non-conscious volition. *Conscious. Cognit.* 33, 196–203. <https://doi.org/10.1016/j.concog.2015.01.002>.
- Schneider, L., Houdayer, E., Bai, O., Hallett, M., 2013. What we think before a voluntary movement. *J. Cognit. Neurosci.* 25 (6), 822–829. https://doi.org/10.1162/jocn_a.00360.
- Schnell, K., Heekeren, K., Daumann, J., Schnell, T., Schnitker, R., Möller-Hartmann, W., Gouzoulis-Mayfrank, E., 2008. Correlation of passivity symptoms and dysfunctional visuomotor action monitoring in psychosis. *Brain : J. Neurol.* 131 (Pt 10), 2783–2797. <https://doi.org/10.1093/brain/awn184>.
- Schultze-Kraft, M., Parés-Pujolràs, E., Matic, K., Haggard, P., Haynes, J.-D., 2020. Preparation and execution of voluntary action both contribute to awareness of intention. *Proc. Biol. Sci.* 287 (1923), 20192928 <https://doi.org/10.1098/rspb.2019.2928>.
- Schurger, A., 2018. Specific relationship between the shape of the readiness potential, subjective decision time, and waiting time predicted by an accumulator model with temporally autocorrelated input noise. *Euro Neuro* 5 (1). <https://doi.org/10.1523/ENEURO.0302-17.2018>.
- Schurger, A., Hu, P., Ben, », Pak, J., Roskies, A.L., 2021. What is the readiness potential? *Trends Cognit. Sci.* 25 (7), 558–570. <https://doi.org/10.1016/j.tics.2021.04.001>.
- Schurger, A., Mylopoulos, M., Rosenthal, D., 2016. Neural antecedents of spontaneous voluntary movement : a new perspective. *Trends Cognit. Sci.* 20 (2), 77–79. <https://doi.org/10.1016/j.tics.2015.11.003>.
- Schurger, A., Sitt, J.D., Dehaene, S., 2012. An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proc. Natl. Acad. Sci. U.S.A.* 109 (42), E2904–E2913. <https://doi.org/10.1073/pnas.1210467109>.
- Searle, J.R., 2001. Free will as a problem in neurobiology. *Philos. Now.* 76 (4), 491–514. <https://doi.org/10.1017/S00318191010000535>.
- Seghezzi, S., Haggard, P., 2022. *Volition And "Free Will"* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/e46pb>.
- Seghezzi, S., Zirone, E., Paulesu, E., Zapparoli, L., 2019. The brain in (willed) action : a meta-analytical comparison of imaging studies on motor intentionality and sense of agency. *Front. Psychol.* 10 <https://doi.org/10.3389/fpsyg.2019.00804>.
- Seth, A.K., Bayne, T., 2022. Theories of consciousness. *Nat. Rev. Neurosci.* 23 (7), 439–452. <https://doi.org/10.1038/s41583-022-00587-4>.
- Shadlen, M., Roskies, A., 2012. The neurobiology of decision-making and responsibility : reconciling mechanism and mindedness. *Front. Neurosci.* 6:56. doi:10.3389/fnins.2012.00056.
- Shibasaki, H., Hallett, M., 2006. What is the Bereitschaftspotential? *Clin. Neurophysiol.: Official Journal of the International Federation of Clinical Neurophysiology* 117 (11), 2341–2356. <https://doi.org/10.1016/j.clinph.2006.04.025>.
- Shields, G.S., 2014. Neuroscience and conscious causation : has neuroscience shown that we cannot control our own actions? *Review of Philosophy and Psychology* 5 (4), 565–582. <https://doi.org/10.1007/s13164-014-0200-9>.

- Sinnott-Armstrong, W., 2022. What is freedom?. In: W. Sinnott-Armstrong, Free Will. Oxford University Press. <https://doi.org/10.1093/oso/9780197572153.003.0004>, pp. 33–40.
- Sinnott-Armstrong, W., Nadel, L., 2011. *Conscious Will and Responsibility : A Tribute to Benjamin Libet*. Oxford University Press.
- Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., Haggard, P., 2004. Altered awareness of voluntary action after damage to the parietal cortex. *Nat. Neurosci.* 7 (1), 80–84. <https://doi.org/10.1038/nn1160>.
- Sohn, Y.H., Kaelin-Lang, A., Hallett, M., 2003. The effect of transcranial magnetic stimulation on movement selection. *J. Neurol. Neurosurg. Psychiatr.* 74 (7), 985–987. <https://doi.org/10.1136/jnnp.74.7.985>.
- Soon, C.S., Brass, M., Heinze, H.-J., Haynes, J.-D., 2008. Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11 (5), 543–545. <https://doi.org/10.1038/nn.2112>.
- Soon, C.S., He, A.H., Bode, S., Haynes, J.-D., 2013. Predicting free choices for abstract intentions. *Proc. Natl. Acad. Sci. USA* 110 (15), 6217–6222. <https://doi.org/10.1073/pnas.1212218110>.
- Spence, S.A., Brooks, D.J., Hirsch, S.R., Liddle, P.F., Meehan, J., Grasby, P.M., 1997. A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain: J. Neurol.* 120 (Pt 11), 1997–2011. <https://doi.org/10.1093/brain/120.11.1997>.
- Thaler, D., Chen, Y.C., Nixon, P.D., Stern, C.E., Passingham, R.E., 1995. The functions of the medial premotor cortex. I. Simple learned movements. *Exp. Brain Res.* 102 (3), 445–460. <https://doi.org/10.1007/BF00230649>.
- Thirugnanasambandam, N., Contreras-Castro, F.G., Hallett, M., 2019. Dual-hemispheric transcranial direct current stimulation (tDCS) over primary motor cortex does not affect movement selection. *PLoS One* 14 (12), e0226103. <https://doi.org/10.1371/journal.pone.0226103>.
- Travers, E., Friedemann, M., Haggard, P., 2021. The Readiness Potential reflects planning-based expectation, not uncertainty, in the timing of action. *Cognit. Neurosci.* 12 (1), 14–27. <https://doi.org/10.1080/17588928.2020.1824176>.
- Travers, E., Haggard, P., 2021. The Readiness Potential reflects the internal source of action, rather than decision uncertainty. *Eur. J. Neurosci.* 53 (5), 1533–1544. <https://doi.org/10.1111/ejn.15063>.
- Travers, E., Khalighinejad, N., Schurger, A., Haggard, P., 2020. Do readiness potentials happen all the time? *Neuroimage* 206, 116286. <https://doi.org/10.1016/j.neuroimage.2019.116286>.
- Ullmann-Margalit, E., Morgenbesser, S., 1977. Picking and choosing. *Soc. Res.* 44 (4), 757–785.
- Vaidya, A.R., Pujara, M.S., Petrides, M., Murray, E.A., Fellows, L.K., 2019. Lesion studies in contemporary neuroscience. *Trends Cognit. Sci.* 23 (8), 653–671. <https://doi.org/10.1016/j.tics.2019.05.009>.
- van Eimeren, T., Wolbers, T., Münchau, A., Büchel, C., Weiller, C., Roman Siebner, H., 2006. Implementation of visuospatial cues in response selection. *Neuroimage* 29 (1), 286–294. <https://doi.org/10.1016/j.neuroimage.2005.07.014>.
- Verbaarschot, C., Farquhar, J., Haselager, P., 2019. Free Wally : where motor intentions meet reason and consequence. *Neuropsychologia* 133, 107156. <https://doi.org/10.1016/j.neuropsychologia.2019.107156>.
- Voon, V., Gallea, C., Hattori, N., Bruno, M., Ekanayake, V., Hallett, M., 2010. The involuntary nature of conversion disorder. *Neurol. Now.* 74 (3), 223–228. <https://doi.org/10.1212/WNL.0b013e3181ca00e9>.
- Walter, S., Heckmann, H.-D., 2003. *Physicalism and Mental Causation : the Metaphysics of Mind and Action*. Imprint Academic.
- Wegner, D.M., Wheatley, T., 1999. Apparent mental causation : sources of the experience of will. *Am. Psychol.* 54, 480–492. <https://doi.org/10.1037/0003-066X.54.7.480>.
- Welniaż, Q., Roze, E., Béranger, B., Méneret, A., Vidailhet, M., Lehericy, S., Pouget, P., Hallett, M., Meunier, S., Galléa, C., 2021. Identification of a brain network underlying the execution of freely chosen movement. *Cereb. Cortex* 32 (1), 216–230. <https://doi.org/10.1093/cercor/bhab204>.
- Wolpert, D., Ghahramani, Z., Jordan, M., 1995. An internal model for sensorimotor integration. *Science* 269 (5232), 1880–1882. <https://doi.org/10.1126/science.7569931>.
- Wolpert, D.M., Ghahramani, Z., 2000. Computational principles of movement neuroscience. *Nat. Neurosci.* 3 (S11), 1212–1217. <https://doi.org/10.1038/81497>.
- Yaffe, G., 2022. How is responsibility related to free will, control, and action? In: Yaffe, G. (Ed.), *Free Will*. Oxford University Press, pp. 119–126. <https://doi.org/10.1093/oso/9780197572153.003.0014>.
- Yazawa, S., Ikeda, A., Kunieda, T., Ohara, S., Mima, T., Nagamine, T., Taki, W., Kimura, J., Hori, T., Shibasaki, H., 2000. Human presupplementary motor area is active before voluntary movement : subdural recording of Bereitschaftspotential from medial frontal cortex. *Exp. Brain Res.* 131 (2), 165–177. <https://doi.org/10.1007/s002219900311>.
- Zapparoli, L., Seghezzi, S., Paulesu, E., 2017. The what, the when, and the whether of intentional action in the brain : a meta-analytical review. *Front. Hum. Neurosci.* 11 <https://doi.org/10.3389/fnhum.2017.00238>.
- Zapparoli, L., Seghezzi, S., Scifo, P., Zerbi, A., Banfi, G., Tettamanti, M., Paulesu, E., 2018. Dissecting the neurofunctional bases of intentional action. *Proc. Natl. Acad. Sci. USA* 115 (28), 7440–7445. <https://doi.org/10.1073/pnas.1718891115>.