An intellectual history of the “Libet experiment”: embedding the neuroscience of free will

Haggard, Patrick [1]

[1] *Institute of Cognitive Neuroscience, University College London* DOI 10.5281/zenodo.13341982

TO CITE

Haggard, P. (2024). An intellectual history of the “Libet experiment”: embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982

PUBLICATION DATE

19/08/2024

ABSTRACT

In 1983, Benjamin Libet and colleagues published a paper that profoundly changed psychological, philosophical and neuroscientific discussions about free will. Briefly, Libet showed that the subjective experience of conscious intention followed the readiness potential (a key neural marker preceding voluntary action), rather than preceding it, as an intuitive dualism might imagine. The experiment itself has been much debated, and often criticized, but its positioning and context within the general history of mind sciences has been neglected. This paper takes a history of ideas approach, and shows that the Libet experiment represents the convergence of two distinctively 19th century traditions: a philosophical interest in volition and conation, and a physiological tradition of measuring timing delays. A discussion of the afterlife of Libet’s result identifies a number of crucial methodological controversies and concerns, to which Libet’s arguments from temporal precedence may be particularly vulnerable. Experimental methods for studying volition are inevitably constrained by the fact that, by definition, the input signals for voluntary action are under the control of the participant, and not, as is usual, under the control of the experimenter. Understanding the broader embedding of the experiment within the scientific study of volition may be useful for guiding future research using the same methods, or indeed alternative methods.

Acknowledgements

Many people have helped with this research. The idea for the paper is a direct result of a series of lunch meetings with Michael Pauen, 2022-2024. I would also like to thank Silvia Seghezzi and Chris Frith for multiple discussions of these ideas over previously years. Tim Bayne and Liad Mudrik strongly influenced the discussion of the relation between volition and modern neuroscience of consciousness. Daniel Dennett and Wolf Singer both generously took time to discuss with me how Libet’s work was originally received. Interactions with Aaron Schurger, Uri Maoz, Tomas Dominik, Alfred Mele and many others have all been invaluable in guiding my thinking about the topic. I am grateful to Tavishi Bhalotia for careful proof-reading and manuscript assistance.

The background research and development of the paper was generously supported by a Reimar-Lüst Prize from the Alexander von Humboldt Foundation and the Thyssen Foundation. The preparation of the final manuscript was generously supported by a writing fellowship from the Institut des Études Avancées, Paris.

This paper was written during a 1-month residence at the Paris Institute for Advanced Study under the "Paris IAS Ideas" program *

Volition as a historic preoccupation of psychology

Psychology as an academic subject and branch of knowledge emerged in the 19th century from the combined traditions of philosophy and physiology. Key questions and theories about how the human mind worked came from the traditions of British Empiricism and Enlightenment philosophy, notably from Kant. The key methods for investigating these questions empirically, and thereby revealing the brain mechanisms underpinning mental processes, came somewhat later with the emerging science of neurophysiology.

A contemporary reader of general psychology journals, for example the well-respected Psychological Science, might well miss these dual origins of the discipline. Modern psychology has developed its own theoretical constructs, methods and controversies that seem distinctive, self-sufficient, and encapsulated. In this article, I want to return to a specific topic in psychology where these dual philosophical and neurophysiological origins remain clearly visible, and still pertinent. This is the topic of voluntary action.

The problem of voluntary action

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Arguably, there are two obvious reasons why scientific thought about the workings of mind might emerge in human culture. One is a first-person reason: to investigate and understand one's own subjective experiences. The second reason is both first-person and third-person: to understand causes and processes behind one's own actions, and the actions of others - i.e., to try to explain behaviour. Not all actions are voluntary. For example, strong stimuli can trigger involuntary reflex responses, and stimuli with predictive value can trigger conditioned responses. Yet, providing a convincing positive definition of exactly why some actions are voluntary remains difficult. Traditionally, psychologists have defined voluntary action in terms of freedom of choice between several possible alternatives. For example, in "simple reactions" the participant is instructed to make just one response when a stimulus occurs. In "choice reactions", the participant must select which of several responses to make according to which of several stimuli has occurred (Frith & Done, 1986) In classical choice reaction tasks, the participant chooses and acts on the basis of the stimulus alone. But outside of the laboratory, they might choose on the basis of partly internal information, or some combination of internal and external information. Examples of internal information could include preferences, motivation, or just picking one response at random (Ullmann-Margalit & Morgenbesser, 1977) the origin of behaviour is, at least partly, self-generated, internal and the result of a specific decision process, the action can be classed as voluntary. The process of internal generation of the information that leads to the action, in turn, can be described as a process of volition.

A detailed exposition of the concepts of voluntary action and volition is beyond the scope of this article, but may be found elsewhere (Passingham et al., 2010). Here, I shall restrict myself to two interdisciplinary comments. First, the psychological definition of volition as self-generated action choice seems close to the traditional philosophical concepts of freedom and autonomy (Berlin, 2002) that are considered distinctive features of the human condition. This means that psychological approaches to volition, like their philosophical counterparts, face the difficult challenge of explaining the causal metaphysics of agent causation: how can a mental choice or decision cause a physical action? Second, a simpler definition of voluntariness comes from the neurophysiological concept of a "voluntary motor system" (Sherrington, 1906). For the neurophysiologist, an action is voluntary if it is the result of neural activity in a specific cortical network that links a range of frontal motor areas to the primary motor cortex.

In this article, I take an intellectual history approach to the history of volition in psychology, focussing particularly on the precursors and sequelae of Libet's well-known

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

experimental investigation of 'free will' published in 1983. I will investigate how the dual-origin account of psychology, as emerging from the convergence of philosophy and neurophysiology, might extend to volition. I will argue that Benjamin Libet's experiments on volition published in 1983 fit naturally with psychological research programmes from a hundred years earlier. I will also argue that the current interest in Libet's work, particularly in modern science of consciousness, can still be read in terms of the same historical controversies and trade-offs.

A tripartite mind

A concept similar to volition was clearly present in the 18th century philosophical tradition from which psychology emerged. However, this concept was given a variety of different names. Kant, in his Critique of Judgement (Kant, 1790/2005) wrote "There are three absolutely irreducible faculties of the mind, namely, knowledge, feeling, and desire". Here knowledge refers primarily to knowledge resulting from perception, while 'desire' refers to the mental state that triggers action, particularly goal-directed action. Indeed, the core concept of a mental state associated with voluntary action, has been identified by many different terms, often differing only slightly in emphasis. The term 'will' might equally be used, and indeed often was. In one of the foundational texts of scientific psychology in English, Ward (Ward, 1886) used the term 'conation' (literally 'trying') to express the same idea. More recently, discussions of 'intention' and 'intentional action' carry similar force: an agent has a distinct goal or purpose, which they progress through their action. I will use the terms volition and voluntary action to capture the same construct.

The first proponent of the tripartite division of mental processes taken up by Kant appears in fact to have been Moses Mendelssohn (Hilgard, 1980). Here, it seems useful to identify why volition is so distinct from perception, and from affect, that an additional third category is needed. Perception and affect are, for both Mendelssohn and Kant, reactive – they are mental processes caused by external events or circumstances. The external event provides the content of the mental state. In contrast, the causal structure of volition appears to be quite different. Our voluntary actions are not merely reactive responses to external events, but have a mind-to-world direction. Many voluntary actions depend on internally generating the mental representation of a counterfactual goal that seems to pull the motor system towards its fulfilment.

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Let me give a simple example of what this could mean. Somebody who is asked to kick from the penalty spot in a soccer match has the goal of kicking the ball into the net. They know that the best way to do this is to kick towards either the top left or the top right of the net. They also know that the goalkeeper should be unable to predict in which direction they will kick. If a fellow player were to shout out “Aim left!”, it would be unwise for the kicker to simply follow this instruction, as the goalkeeper may also have heard it. They should thus generate purely internally a decision about which way to kick, and this decision should determine their action. Clearly, many other cognitive and contextual elements are also required for the action to succeed: the player must see where the ball is, and where the net is, they must retrieve from memory their knowledge of how to kick the ball, they must sense and react to external factors like the wind. However, the distinctive core of the action is an internally-represented goal and an internally-generated means of achieving it – in t. The action is therefore considered voluntary because it is marked by the lack of external inputs, particularly whether to kick right or left. Interestingly, similar internal generation of ends and means is also associated with other distinctly different areas of psychology, such as creativity and problem-solving.

Mental Chronometry

As psychology successfully transitioned towards an experimental science in the late 19th century, experimental study of volition was neglected. Significant progress was made in understanding the mechanisms of perception, but scientific investigations of volition were scarce. In many ways, the starting point of scientific psychology was Helmholtz’s realisation that the operations of the nervous system took measurable time. From this, it follows that the measurement of elapsed time can offer important information about the nature, and particularly the complexity of mental operations. Wilhelm Wundt, Helmholtz’s student and assistant, focussed on estimating and interpreting the time of subjective events. Strongly inspired by Leibniz’s concept of apperception, Wundt saw the timing of perceptual representations as an important window into the mechanisms of conscious experience. Over a period of years, Wundt and his students developed an extensive research programme of mental chronometry that provided perhaps the first systematic approach to studying the mind experimentally. Interestingly, this scientific development largely depended on a technical development that Wundt had mastered during his work with Helmholtz: the Hipp chronoscope. This was a customised clock

Haggard, P. (2024). An intellectual history of the “Libet experiment”: embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

mechanism that could measure, with millisecond accuracy, the delay between a starting event and a stop event.

The chronoscope made rigorous studies of reaction times and perceptual latencies possible – as long as convincing start and stop events could be identified. A detailed summary of these experiments is also available (Wontorra et al., 2011). Wundt also developed a 'complication experiment' which dispensed with the motor response. In this experiment, participants observed a moving visual stimulus (originally a pendulum, but in later versions a clock hand), and noted the position of the moving stimulus at which they perceived a second stimulus, such as the onset of a sound. Wundt referred to the additional, target stimulus as a "complicating" stimulus. This perceptual chronometry approach therefore involved judgements of cross-modal synchronisation between the moving visual stimulus and the complicating stimulus. This method forms the basis of Libet's experiment.

Wundt himself made important theoretical contributions to the study of will (Fahrenberg, 2019), but the mental chronometry experiments in his laboratory appear to have all been perceptual in nature. In perception, the experimenter triggers the onset of a stimulus, which marks the input to the relevant mental processing module, and provides the start event for the chronoscope. Interestingly, Wundt believed that simply attending to a sensory input, which he called apperception, was in fact a volitional mental act. But if we instead consider volition as the endogenous generation of actions, then Wundt's perceptual approach to mental chronometry faces an irresolvable problem: the starting events for a voluntary action are, by definition, internal to the agent themselves, and not available as the starting signal for any chronoscope.

Wundt's research programme gave rise to two distinct traditions in experimental psychology (Boring, 1950). One tradition stayed close to chronometric stimulus presentation and response measurement, and used these paradigms to search for lawlike regularities in perception, attention memory, and other mental processes (Titchener, 1908M). Oswald Külpe, Wundt's assistant for several years, founded a second tradition by extending these subjective report methods from perception to thought and higher cognition. This approach became associated with the Würzburg School, and ultimately provided the basis for Gestalt psychology. Interestingly, some of the psychologists within this second tradition definitely did try to tackle volition – but in doing so, they subtly shifted the emphasis, and finessed the problem of the missing input signal.

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

To understand this, we need to return to the definition of volition. Up to now, I have presented volition simply as the mental state that internally generates goal-directed actions. In fact, volition is a more multifaceted and even elusive concept. Many properties can be associated with voluntary action. No single one of these may be necessary for an action to be voluntary, but some combination of them may be jointly sufficient. Some of the properties that make an action voluntary seem not to be direct properties of the action itself, but rather counterfactual properties related to the agent and the context. Trying to characterise precisely what makes an action voluntary produces at least the following elements:

Table 1. Six Features of Voluntary Action

Feature
The action is a motoric expression of a cognitive intention
The action is endogenous, in the sense of independent of current sensory stimulation
The action is goal-directed, reasons-responsive, motivated
The action is spontaneous
The agent could not have made the action that they made
The agent suppresses actions that they did not want to make

Perhaps unsurprisingly, no single experiment can operationalize all of these factors simultaneously. Most experiments on volition, since the very earliest days, emphasized one factor while ignoring or compromising on others.

The missing mental chronometry of volition

Michotte and Prüm (1910) devised an ingenious experiment to elicit voluntary actions, working within the introspectionist traditions of the Würzburg school. They briefly presented two numbers, for example 65 and 43, or 54 and 16. The participants viewed the two numbers, and decided for themselves whether to add them or subtract them. In some versions of the experiment, the participants then reported the result of the sum, in other versions they merely closed a chronoscope stop-switch once they had decided. Interestingly, essentially the same task was used almost a hundred years later in a highly-cited publication on decoding voluntary choices from fMRI signals in the human brain (Soon et al., 2008). Michotte and Prüm collected introspective reports from a small group of trained observers, and the qualitative part of their work was perhaps

Haggard, P. (2024). An intellectual history of the “Libet experiment”: embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

more accomplished than their quantitative observations. They noted experiences of anticipatory tension, of insight at the point of decision, and of a distinctively endogenous quality to this decision. However, irrespective of how endogenous the actions might feel, one might question how endogenous they really are: 65, 43 invites subtraction while 54, 16 invites addition. These apparently endogenous actions in fact depend strongly on rules of arithmetic that have been overlearned and stored in memory, coupled with a general principle of minimizing cognitive effort. This approach therefore seems to particularly miss the second and fourth aspects of volition in the above table.

Narziss Ach (Ach, 1905) took an interestingly different approach. He first required participants to thoroughly learn a list of paired associate words. He then gave them the instruction to produce a rhyme as quickly as possible after visual presentation of a given word. The participant was considered to develop a determiniende tendenz, essentially a contingent volition, due to this instruction. For Ach, the key measurements occurred when a word from the paired associate list was later presented as a stimulus in the rhyming task. To produce a novel word that rhymes with the presented word, the participant has to resist the overlearned, prepotent tendency to respond with the word that was previously associated with the presented word. The speed and accuracy with which they produce a suitable rhyme word were taken as an index of the strength of will engendered by the instruction to produce rhymes.

Ach's work is interesting in several ways. First, volition is, once again, not truly endogenous, since it is given by an experimenter instruction ("produce a rhyme word"). One might now call it a 'task set' (Monsell, 2003) or behavioural rule, rather than an intrinsic generative process leading to desired action. Second, volition figures as a prior condition before the task is executed, and therefore out of scope for mental chronometry. Third, and most importantly, Ach's experimental approach gives an early example of the intertwining between what I have previously called positive and negative volition (Seghezzi & Haggard, 2023). Positive volition refers to the generation of action, while negative volition refers to inhibit inappropriate or unwanted actions. Negative volition is clearly linked to inhibitory cognitive control, and to the folk-psychological notion of "willpower", in the sense of resistance to temptation (Mischel & Ebbesen, 1970). In Ach's task, generating a rhyme word would involve positive volition, while inhibiting the tendency to respond with the previously-learned associate would involve negative volition. Scientific discussions of volition, and philosophical discussions of 'free will' both regularly conflate positive and negative volition. This may seem surprising, since positive and negative volition have clearly opposite behavioural signs,

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

namely action and inaction respectively. Focussing on negative volition has several practical advantages for the experimenter, as Ach showed. It bypasses difficult causal questions about the starting triggers for volition; it gives the experimenter the familiar and powerful tool of stimulus control, and it gives the powerful measurement tool of studying a cognitive system through its failures, since any prepotent response to the stimulus can be interpreted as a failure of negative volition.

Ach's approach therefore stresses aspects 5 and 6 from table 1. It again seems deficient on aspects 2 and 4, though for different reasons from Michotte's work. Thus, neither Michotte nor Ach bring into clear focus the first and third aspects of volition from table 1, which perhaps come closest to the traditional philosophical constructs of freedom and autonomy. Thus, the nascent experimental psychology of volition failed to engage with the familiar idea that the healthy adult human mind is somehow involved in our capacity to act in the way we desire, to achieve the goals we have given ourselves.

The readiness potential

In 1965 Kornhuber and Deecke reported a newly-discovered brain activity, which they called the readiness potential (Kornhuber & Deecke, 1965). They asked participants to make occasional self-paced hand movements, while they recorded EEG continuously. By averaging the EEG records time-locked to the moment of action, they were able to reveal a gradual ramp-like increase in negativity, lasting a second or more, and peaking at the moment of voluntary action. Their discovery had an important transformative influence on the field, since it seemed to identify the precursors of voluntary action. They were cautious not to claim discovery of the elusive endogenous input signal for volition: after all, EEG signals are but coarse and incomplete measurements of neuronal activity, and may not catch the true physiological input signal. However, they did view the readiness potential as a biomarker of volition. If that is true, then the onset of the readiness potential becomes at least a key proxy for the command signal for volition: this was certainly the way it was later interpreted by Libet, as we shall see.

The significance of the readiness potential has been extensively debated in recent years (Shibasaki & Hallett, 2006). and a full review is beyond our current scope. In an elegant and influential study, Schurger and colleagues (Schurger et al., 2012) showed that the readiness potential could simply be an artefact of averaging stochastic noise on

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

which participants might rely for triggering actions in the somewhat unusual experimental situation of freely voluntary acts. We return to this critique later.

What did Libet do?

Benjamin Libet had a long and distinguished career in neuroscience research. His book Mind Time (Libet, 2004) contains elements of scientific autobiography as well as a detailed review of his work on consciousness, time experience and the brain. For many decades, Libet's main research interest lay in human somatosensory neurophysiology. By direct stimulation of the somatosensory cortex as part of clinical neurosurgical procedures, he showed that sensory stimuli are consciously perceived only if they involve cortical activities sustained for a period of some 500 ms (Libet et al., 1979). Direct stimulation of the cortex for shorter periods produced no experience at all. Very brief skin stimuli close to threshold intensity were perceived if they lead to cortical activation lasting 500 ms, but not otherwise. Essentially, consciousness seemed to be the output of a cortical accumulator of neural activity, which included a substantial integration time constant. Libet's distinctive contribution was to consider how people are able to experience stimulus onsets, even though experience is the result of long accumulation. He hypothesised that any afferent sensory volley lays down a thalamic marker as it ascends the somatosensory pathway. If the cortex accumulates sufficient activity as a result of the volley, the marker is read by a backward referral process presumably involving corticothalamic feedback. The time of the marker is then taken as the subjective time of the event.

Libet's theory of sensory timing was perhaps the first neuroscientific account of conscious events, as opposed to conscious contents. That is, it gave a neuromechanistic account of when people have experiences. This supplements the more familiar question of what they experience, and in fact provides a valuable, robust and implicit method for approaching this question. Libet's model convincingly explains why the perceived onset of a stimulus may not coincide with its physical onset (Morton et al., 1976). It thus provided a neural model for the psychological phenomena that Wundt had highlighted a century or so before. Finally, it still provides one of the few models of the interplay between cortical and subcortical factors in consciousness. Unfortunately, the early passing of Libet's neurosurgical colleague brought an end to this unique series of studies.

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

The impact of Libet's sensory studies was amplified by Daniel Dennett's discussion of them (Dennett, 1991). Dennett used Libet's findings to argue against 'Cartesian theatre' views of perception, and in favour of a 'multiple cuts' or 'multiple drafts' models. In these models, consciousness represents a potentially nonlinear rearrangement of widely distributed neural activity, as the brain attempts to integrate multiple inputs to provide a coherent structure of experience. The process of active rearrangement and editing of signals indicates that, for Dennett, consciousness is constructed, rather than passively received.

Libet reported having the idea for his experiments on volition in 1977 (Libet, 2004). The key paper was published in Brain six years later. The intellectual continuity between the volition experiments and Libet's earlier sensory experiments is striking. Libet was clearly aware of Kornhuber and Deecke's work, and was inclined, like them to treat the onset of the RP as marking a specific physiological event, which could be considered the internal cause of voluntary actions. Further, he interpreted the rising RP as a neural accumulation, analogous to those that caused "neuronal adequacy" when delivered to sensory cortex. The key remaining question became when the conscious experience of volition is perceived, relative to these cortical activities. Does voluntary action involve an analog of the thalamic marker that Libet had identified for sensory stimulation, tying consciousness to the start of the cortical activity, or does consciousness enter only at the end of cortical accumulation? The question seems to follow naturally from the convergence of Libet's earlier work on sensory experience, and his awareness of Kornhuber and Deecke's work on the readiness potential.

To address this question, Libet used a method that recalls Wundt's mental chronometry programme, and particularly the "Komplikationspendl" experiments in which participants used a clock to report the moment at which they registered a sensory event (see above). Libet's method and results have been reviewed and critiqued many times before, with one particularly comprehensive review appearing very recently (Dominik et al., 2024).

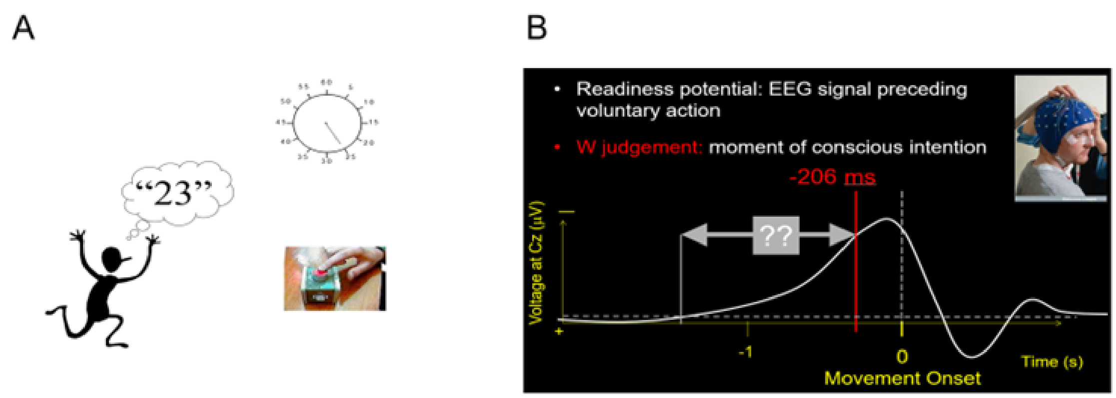
Therefore, the discussion here is limited to what is needed to understand the convergence of ideas that his experiment represents (see figure 1). Libet asked his participants to make occasional voluntary movements, at a time that they chose – effectively, this is a version of the Kornhuber and Deecke experiment. While doing so, they watched a rotating clock hand. They noted the position on the clock hand at which they "first felt the urge to move" – effectively, this is a version of Wundt's "complication" experiments in mental chronometry. The clock stopped at a random delay

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

after the movement, and participants reported the corresponding clock judgement. Libet labelled this time W, since he considered it a judgement of conscious will.

Libet's core result is deceptively simple. The onset of the RP in his data occurred around 750 to 350 ms prior to movement onset, while the average time of the W judgement was around 200 ms prior to movement onset. This temporal ordering seems unsurprising, even inevitable, from a materialist neuroscientific perspective that views conscious states as consequences of brain activity. However, it makes a start contrast with the everyday dualistic intuition that conscious thought somehow causes our actions – if that were true, then the W judgement should precede the RP, yet the opposite is the case. Since this dualistic intuition that conscious intention causes action appears to the basis of legal doctrines of responsibility and mens rea, the question cannot simply be dismissed as an academic nicety. The subtitle of Libet's experiment provides a clear summary of his findings: "The unconscious initiation of a freely voluntary act".

Figure 1



- *Figure 1A. The basic setup of the Libet experiment. Participants view a rotating clock hand, and make a voluntary action at a time of their own choosing. The clock then stops in a random location, and participants indicate the position of the clock hand at which they "first felt the urge to move".*

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

- *Figure 1B. Schematic of Libet's findings. A ramp-like increase in EEG negativity (readiness potential) preceded the time of voluntary action. In EEG averages, the RP can appear to begin over 1 s prior to action. The average time of the "conscious urge to move", or W judgement, occurred much later, and only just before action onset. The temporal precedence and gap between brain activity and conscious intention (grey arrow) appears inconsistent with the dualist intuitions of 'conscious free will'.*

The standpoint in this paper is that of history of ideas. From this perspective, it seems that Libet's experiment lies at the conjunction of three traditions that had each been aired individually one hundred years before, but had somehow failed to converge into specific scientific projects until Libet's experiment. First, Helmholtz's concept that neuronal conduction took measurable time readily explains the finding that consciousness depends on an accumulation of neural activity, in motor just as in sensory cortices. Second, Libet's W judgement clearly recalls Wundt's view of conscious experience as a mental action that takes place at a certain timepoint, and therefore fits with the broad research programme of mental chronometry. In this context it seems interesting that Libet did not cite Wundt directly, though his attention may have been brought to Wundt's earlier work by peer commentary (Breitmeyer, 1985). Finally, Libet's work invokes the concept of a neuronal go-signal that initiates voluntary action, as hypothesized by Kornhuber and Deecke, and as prefigured by discussions of conation going back to Mendelssohn and continuing throughout the 19th and early 20th centuries. In some ways, therefore, Libet's experiment seems more continuous with the theoretical traditions of 1883 than with the neuroscientific traditions of 1983.

On the other hand, one crucial enabling methodological ingredient of Libet's experiment was the capacity to record from the volitional brain before movement, and this did not become available until the discovery of the RP. The RP appeared to make up for the historical lack of any experimentally accessible input signal to volition, and made the project of investigating volition seem empirically feasible for the first time.

The afterlife of the Libet experiment

In what follows I shall critically evaluate some sequelae of Libet's experiment. Debate about Libet's results was immediate, and has been sustained since. This review does not aim to be comprehensive. A wide set of critical reactions followed Libet's 1985 target

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

article (Libet, 1985). A brief summary of some common criticisms of Libet's work is given in the table below.

Table 2. Some critiques of the Libet experiment as an investigation of volition

Critique
The experimental instructions are bizarre and equivalent to a contradiction: "Be voluntary!"
The actions are meaningless and lack any convincing context of volition as it occurs in everyday life
The sampling of both awareness and brain activity are biased by being made contingent on having just made an action
The phenomenal experience of the "urge to move" is weak and elusive
The reported experience could be a retrospective inference or confabulation, unrelated to the actual mechanisms of action generation

Recent reviews have provided comprehensive critical re-appraisal of Libet's work (Dominik et al., 2024). A particularly busy debate in philosophy has considered how Libet's result impacts compatibilism (the idea that human free will may be compatible with the deterministic nature of the universe). This debate revolves, in my view, around what is meant by the concepts of freedom and free action. Is the picture of voluntary action that emerges from the Libet experiment sufficient for us to think of human action as truly free? This question is clearly important, but it is not primarily a psychological question about the voluntary action mechanisms in the human brain, and I will not discuss it further here.

Instead, I will consider three questions about the afterlife of Libet's experiment. Is Libet's result valid in the light of subsequent research? What were the effects of the result on Libet's own thinking? What have been the effects on the broad field of mind science, some forty years after the original publication?

Is the result valid?

As we have seen, the core result concerns the temporal order of three events: first the onset of brain preparation indexed by the RP, then an experience of conscious intention, and finally the motor action itself. Of these three events, only the action itself is clear, unambiguous and ontologically secure. The RP and the moment of conscious intention remain controversial both in terms of their ontological status, and in terms of their

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

measured values. If the events themselves are disputable, then the knowledge gained about their causal relations by establishing their temporal order is also disputable.

A major controversy has recently arisen regarding the interpretation of the RP. Libet followed Kornhuber and Deecke in viewing the RP as a signal corresponding to the cause of voluntary action. On that view, the start of the RP may be considered to measure an input signal to a neural system specific to volition. This logic turns out to be flawed for two distinct reasons: the RP may not be a reliable index of volition, and the RP might not actually represent a consistent starting event. The underlying arguments for these two views are quite technical, but have high scientific importance.

First, the RP is identified by a process of biased sampling, which involves retrospectively measuring the EEG once an action has occurred. It remains possible that RPs in fact also happen at other times, when no action is present. These other RPs would not be seen, because only action-locked events are analysed in an RP experiment. If RPs indeed happen without actions, it would be unwise to treat the RP as a reliable biomarker of volition. Identifying whether RPs do or do not happen in the absence of voluntary action is not straightforward (Travers et al., 2020), and this question remains unresolved. The poor signal-to-noise ratio of the RP means that they are not readily detectable by simple inspection. The standard method of improving signal-to-noise by averaging, which allows an average RP to be seen despite high levels of noise on any individual trial, requires an event to define the epochs which will be averaged, and we have no particular hypothesis about what event should be used to identify those hypothetical RPs that are not related to action.

Schurger et al. (Schurger et al., 2012) have suggested that the RP may, in fact, be only stochastic noise, without any signal in the conventional sense. The authors hypothesised that participants might use the accumulation of an internal neural noise to trigger movement in the unusual situation of the Libet experiment. If several recordings of the noise are then aligned to the moment of action, it follows necessarily that the average of the recordings will show a gradual increase towards the threshold level that triggers the action. This would reproduce the form of the average RP, even though no RP-like signal need be present on any individual trial. The RP could, therefore, not be a true signal at all, but an artefact of averaging noise. The noise is systematically related to action, because, according to the model, it in fact triggers action. In this case, the apparent start of the RP is not a specific neural event, and should not be reified or interpreted as the onset of any specific process, such as volition.

Haggard, P. (2024). An intellectual history of the “Libet experiment”: embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Interestingly, the illusion that there is a distinct neural event at the start of the RP may be compounded by the traditional processing pipeline used to calculate the RP. Individual EEG epochs are typically baseline-corrected, to adjust for slow shifts in voltage unrelated to the process of interest, and to reveal only those events that are related to the process of interest. For stimulus-locked EEG, the baseline is typically set just prior to the stimulus. However, this approach is not available for RP, since there is no stimulus. RP researchers have instead generally defined an arbitrary premovement baseline period, and used this for baseline correction. However, this approach carries the clear assumption that the process of volition occurs only after the baseline period, and not during or before it. In fact, we of course do not know when volition begins: that is the question under investigation. Therefore we do not know when to select a baseline when analysing the RP. Moreover, baseline correction inevitably has the effect of eliminating variability during the baseline period, so that the average signal appears to grow from nothing immediately thereafter – this might be mistaken for the onset of an event, but is of course an illusory artefact of the processing pipeline. We have proposed (Khalighinejad et al., 2018) performing baseline correction at the time of action, rather than at some arbitrary-selected premovement baseline period. This may help to avoid confounding the end of the baseline period with an onset of a precursor process.

Schurger et al.'s model is a competence model, rather than a performance model. In essence, it shows that an action-triggering stochastic noise accumulator could generate an RP, but it cannot conclusively show that this is indeed the source of actual RPs. To test this second hypothesis, one could potentially investigate RP duration across single trials. The classical view would predict a consistent duration of the RP, corresponding to the duration of the underlying cognitive process. The stochastic view would predict that RP duration is highly variable, with the precise distribution depending on the details of the underlying stochastic noise. Dirnberger and colleagues (Dirnberger et al., 2008) performed one of the very few studies to have investigated whether the RP does indeed have a consistent onset time, and they concluded that it did.

Any time-series can be decomposed into a signal plus noise. In a sense, the classical RP view and the stochastic noise view can be seen as opposite ends of a continuum extending from a high to a low signal-to-noise (SNR) ratio. Interestingly, Schurger et al.'s model includes an initial baseline shift that bring the noise close to the action threshold – this shift might be considered an implicit signal that accompanies the noise. A recent simulation study (Bogler et al., 2023) suggested that the contribution of stochastic noise might in fact be limited, and that the baseline shift plays an essential

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

role even in a stochastic model of the Libet experiment. This could be an important area for future research.

The second key event in Libet's work is the W judgement itself: the moment of conscious intention. How valid is Libet's estimate of -206 ms for the appearance of a conscious 'urge to move'? The precision of cross-modal synchronisation judgements, such as between an internal signal and a rotating clock, is low. In particular, the 'prior entry' effect (Spence & Parise, 2010) means that whether one attends more to the clock or more to one's own internal signals will introduce a systematic bias in synchrony judgements. Specifically, events on an attended stream are perceived to occur earlier than simultaneous events on an unattended stream. In fact, this bias is small relative to the temporal gaps separating the three events (RP onset, W judgement, movement onset) that Libet studied, so it is unlikely to overturn the temporal sequencing that he reported.

A more pressing difficulty in interpreting Libet's W judgements is to understand what participants are reporting. The phenomenology of voluntary action is often thin and indistinct: one just moves. One may know and feel that an action is fully voluntary, yet lack any clear feeling of a moment of "urge". Helmholtz was perhaps the first to appreciate that many percepts involve a strong element of inference. In the case of the Libet experiment, participants report the time of volition only after they have acted. They might therefore retrospectively infer from the fact of having acted that there must have been a preceding process of volition. On this view, the W judgement would not reflect the perception of any subjective event, but simply beliefs about when one is likely to decide about their actions. This retrospectivist interpretation was strongly argued by Daniel Wegner (Wegner, 2002), who suggested that conscious will was effectively an illusory narrative or retrospective confabulation, rather than a genuine mental state. People might have no conscious experience at all corresponding to their impending voluntary actions, but might confabulate or infer such experiences after the fact, in order to create a coherent narrative around the actions they have just made.

The retrospectivist inference position appears not to have been considered by Libet himself, even though his earlier concept of thalamic markers involved a form of retrospective tagging. Retrospective inference of the kind proposed by Wegner doubtless occurs, and may form part of the normal stream of consciousness, with its multiple redraftings. However, a wholly retrospectivist account of conscious volition is unconvincing. If people are interrupted at random with the question "Are you about to act?", then they are able to answer the question, suggesting some prospective experience

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

of volition. The fact that neural precursors such as RP and beta-band desynchronization (Parés-Pujolràs et al., 2023) are found to precede the question if the answer is affirmative, but do so to a lesser extent if the answer is negative, suggests that such prospective volition is not purely confabulatory, but has its bases in the cortical activities that precede and perhaps produce movement. An ingenious experiment developed by Matsuhashi and Hallett (Matsuhashi & Hallett, 2008) effectively measured the prospective experience of volition, by randomly probing people while they occasionally made voluntary actions. Their results suggested that the timing of conscious intention is much earlier than Libet's W judgement, at around 1200 ms before the expected onset of the impending movement. One might argue that this estimate precedes the typical onset of the RP. But arguments about temporal precedence involving RP seem to me fraught and uncertain: we do not know whether the RP truly has an onset, and we do not have an unbiased estimate of when it may be.

Influence of Libet's result on his own theory of volition

Libet believed that the temporal order of brain events and subjective events ruled out the traditional view that voluntary actions are caused by a process of conscious free will. However, he believed that there was sufficient time, around 200 ms, after the moment of conscious intention to veto an action that one was about to make. This hypothesis was suggested in the 1983 paper, though no direct evidence for a veto process was offered. Thus, although we may lack conscious free will, Libet suggests that we may nevertheless retain the possibility of conscious free won't. Importantly, this means that there is still a real sense in which our actions are up to us, and we therefore become responsible for them. A large part of his book Mind Time is devoted to defending the conscious veto hypothesis.

Historically, Libet's argument recapitulates a move from positive volition (internal generation of action) to negative volition (the capacity to veto unwanted actions) that has been widespread in the history of volition. We have already seen that Ach's experimental method for studying the will involved a similar move. A recent focus on the concept of willpower in social psychology made the same move (Baumeister & Tierney, 2011), but see also (Vohs et al., 2021). Interestingly, Schurger et al.'s work is often seen as consistent with a form of veto. While strongly arguing against an early

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

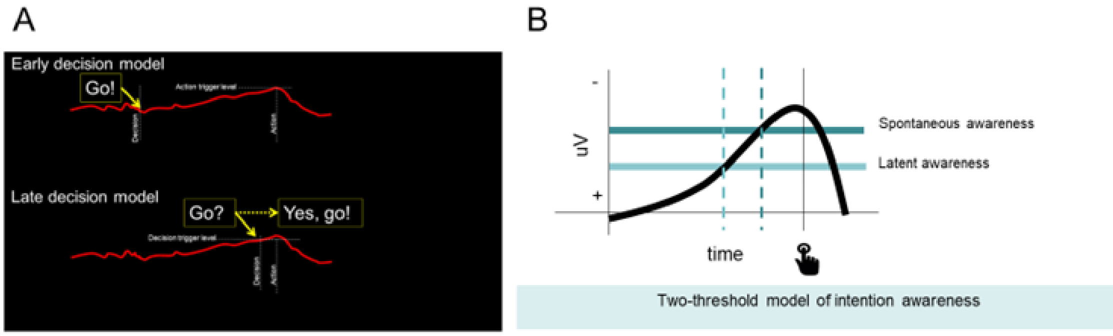
decision that initiates the process of voluntary action, Schurger et al.'s finding is often seen as consistent with a late decision, made after stochastic noise breaches the action-triggering threshold, about whether to go through with the action or not. Stop-signal tests suggest that a veto or countermanding process may occur up to some point of no return (Logan, 2015).

Modern neuroscience of voluntary action has tended to conflate positive and negative volition. I speculate this is because they have similar socio-ethical consequences: either positive or negative volition would be sufficient for holding agents responsible for what they do. But, positive and negative volition are necessarily very different processes from a functional point of view: one promotes action, while the other prevents it. Positive volition recalls the classical concepts of conation and desire, while negative volition corresponds to the neuropsychological concept of inhibitory control (Aron, 2007). Negative volition is more amenable than positive volition to study in controlled laboratory conditions, as Ach showed. From the point of view of neurophysiological signals, positive volition locates executive control and decision at the start of the preparatory precursor processes of movement preparation, while negative volition locates it at the end of these processes, at the moment of a final action trigger (see figure 2).

In future research, improved methods for eliciting and analysing the precursor signals that precede voluntary action could play an important role in clearly understanding the respective roles and neural correlates of positive and negative volition. The early/late controversy for volition has interesting similarities to a controversy surrounding early vs late selection in selective attention (Makovski et al., 2014). We have already seen that, for Wundt at least, volition and attention were similar mental states. Detailed temporal analysis of the time course of selective attention is now possible (Hosseini et al., 2024). Future research might potentially use similar methods to investigate volition.

Figure 2

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.



- *Figure 2A: According to early decision models, an early decision to act causes the start of neural preparation for action, of which RP is a measure. This preparation continues until a trigger level for action is reached. According to late decision models, in contrast, an identified neural signal fluctuates, and occasionally reaches a trigger level, resulting in a decision whether to procede to action, or not.*
- *Figure 2B: People may report that they are about to act if they are probed with a question if some relevant neural activity (here schematised as an RP-like quantity) exceeds some low 'latent awareness' threshold. In contrast, the awareness of intention will not come to dominate experience until a higher threshold is reached, leading to a spontaneous awareness of 'urge to move'. The difference between the two threshold levels for awareness is primarily a matter of attention. Note that drawing attention to intention has an effect equivalent to favouring early decision models over late decision models.*

Libet's concept of conscious veto has been received cautiously in neuroscience because Libet took an explicitly dualistic view of the veto process. Moreover, he believed that veto could involve a causal process in which a purely conscious field could act on neural processing in the cortex to prevent the dispatch of a motor command. Here it becomes interesting to relate Libet's dualism to the mind-brain interaction views of John Eccles. It is clear from Libet's book Mind Time that he was substantially influenced by Eccles' thought. Eccles held that purely mental processes could influence physical processes in the brain, via a quantum interaction. He may have conflated the quantal release of synaptic neurotransmitters with the concept of a quantum probability field, despite the clear differences between these phenomena. In the specific case of voluntary action, he held that the mind could capitalize on moments of high excitation in the supplementary motor area, to increase the probability of synaptic exocytosis, and thus

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

trigger movement (Beck & Eccles, 1992). While Eccles' explicit dualism found few supporters, the core elements in his system, namely the special role of the medial frontal cortex, and the importance of fluctuations in neural 'noise, are clearly present in Libet's work and the work of others in the field. Interestingly, Libet attended a study week organised by Eccles at the Pontifical University in 1964. Libet's paper on that occasion dealt with cortical mechanisms of somatosensory experience. We cannot know the details of the meetings and discussions that Libet had with Eccles, but it seems possible that the core theoretical interest in the causal relations of mind and brain during voluntary action emerged either at this meeting or in later follow-up meetings. Interestingly, the meeting summary noted that the scientific understanding of brain mechanisms of voluntary action lagged far behind the understanding of perception – again reflecting the lack of experimentally-accessible input parameters. Libet reports that the idea for the 'clock method' that enabled his 1983 experiment had first occurred to him in 1977 (Libet, 2004), but his earlier sensory work already makes clear his interest in mind-brain interactionism. It seems likely that his interests in 'free will' were strongly influenced by his discussions with Eccles.

The influence of the Libet experiment on mind sciences

I have already mentioned the important influence of Libet's experiment on the philosophy of free will, without discussing this in detail. Within cognitive neuroscience, there have been attempts to replicate the experiment, and critiques of its physiological and psychological constructs, which have been reviewed extensively above, and also elsewhere (Dominik et al., 2024). Here I want to focus, instead, on the reception of Libet's experiment in modern consciousness science.

Understanding the mechanisms of consciousness is widely seen as a key challenge for modern science. Empirical projects in consciousness science divide into those that focus on the level or state of conscious mentation as an attribute or state of an individual brain (Laureys & Schiff, 2012) and those that focus on experience: i.e., phenomenal content linked to a specific external event. Libet's work speaks to the second tradition, since it deals with the experience of a specific impending action. The dominant models of conscious experience tend to be based on visual experience, and they emphasize the integrating, coherent nature of consciousness (Bayne et al., 2024). All of them have a

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

non-trivial resemblance to Wundt's concept of apperception, and indeed to current concepts of working memory – what I am attending to or thinking about now. These theories have the goal of explaining how what I see can, at least sometimes, become so salient and so meaningful for me. In global neuronal workspace theory (Mashour et al., 2020), consciousness corresponds to the simultaneous rebroadcasting of a specific input content to multiple brain networks. In Dennett's multiple drafts model, conscious experience is a mixed amalgam of multiple inputs, typically combined in a nonlinear way, in an effort after meaning.

These are therefore models of how inputs are centrally processed, and tell us rather little about the conscious experience of outputs, such as the command signals for voluntary action. One simple theory might be that we know our action only as inputs through sensory feedback generated as our body moves. This feedback would trigger a mental reconstruction of the intention that caused the action. Perhaps the efferent motor command plays the same role in backdating the intention that Libet's proposed thalamic marker played in his somatosensory studies. However, extending these theories of consciousness to volition seems to me to meet two fatal obstacles. Global neuronal workspace theories cannot readily explain why the experience of volition is so often like a background buzz in mental life, rather than a total dominance of all current thought. Multiple drafts/cuts models cannot explain why an experience of volition can arise prospectively, without any specific inputs needing to be reprioritized and rearranged.

In IIT, conscious free will is viewed as intrinsic power of an appropriately organized integrated system. This power arises when from the capacity of highly-integrated brains, and perhaps other highly-integrated systems, to envisage alternatives, and to change the world accordingly (Tononi et al., 2023). At this point, volition becomes less like a physiological command signal, and more like the capacity to represent counterfactual goals that was described above. Libet saw a deep link between conscious volition and the voluntary motor command, while current theories of consciousness are generally input-oriented rather than command-oriented. It has been suggested that conscious volition is a distinct perception-like signal (Matsuhashi & Hallett, 2008), but this view has not yet particularly impacted current theories of consciousness.

Conclusion

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Libet's experiment was published over 40 years ago, yet remains a live topic in scientific and scholarly debate. It is often criticized, or even ridiculed. However, by taking a history of ideas approach, I hope to have shown that Libet was perhaps the first person to make positive volition a topic for experimental investigation. He devised and achieved an experiment that would have been suggested by the ideas circulating at the very earliest phases of the discipline of scientific psychology, but which did not actually happen until a century later. In that sense, Libet set out to answer a foundational question that had been posed long before, but had not been satisfactorily addressed by previous scientific investigations. That question is: how and why do people feel that they are the author of their own actions. Perhaps, Libet offered a partial answer to this question. The uncontroversial part of his answer is the physicalist conclusion that our volition, and our conscious experience of it, is a brain mechanism. How that mechanism works remains controversial, and modern research has questioned Libet's methods, revealing a number of methodological confounds. The correct interpretation of the Readiness Potential remains an important topic of current research and controversy. Whether the flaws in the Libet experiment are fatal to its physicalist conclusions remains, as yet, unclear.

A second controversial element of the experiment is Libet's own suggestion of a veto mechanism. He offered no convincing direct evidence for this mechanism, but attributed an implausible residual dualism to it. The veto aspect continues to dominate many current discussions of the experiment. In my view, the field should clearly distinguish Libet's empirical contributions to positive volition from his theoretical suppositions about negative volition.

Libet's experiment remains important in psychological thinking today, and has recently become an important focus in the scientific study of consciousness. However, many discussions of the meaning of the Libet experiment for understanding consciousness suffer from taking an input-based perspective on human cognition. Conation still remains an elusive third pillar in studying the mind.

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Bibliography

Aach, N. (1907). *Uber Die Willenstatigkeit Und Das Denken*. Vandenhoeck & Ruprecht.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *13*(3), 214–228. https://doi.org/10.1177/1073858407299288

Baumeister, R. F., & Tierney, J. (2011). *Willpower: Rediscovering the greatest human strength*. Penguin Press.

Bayne, T., Seth, A. K., Massimini, M., Shepherd, J., Cleeremans, A., Fleming, S. M., Malach, R., Mattingley, J. B., Menon, D. K., Owen, A. M., Peters, M. A. K., Razi, A., & Mudrik, L. (2024). Tests for consciousness in humans and beyond. *Trends in Cognitive Sciences*, *28*(5), 454–466. https://doi.org/10.1016/j.tics.2024.01.010

Berlin, I. (2002). *Two Concepts of Liberty*. Oxford University Press.

Bogler, C., Grujičić, B., & Haynes, J.-D. (2023). Clarifying the nature of stochastic fluctuations and accumulation processes in spontaneous movements. *Frontiers in Psychology*, *14*, 1271180. https://doi.org/10.3389/fpsyg.2023.1271180

Bogler, C., Grujičić, B., & Haynes, J.-D. (2023). Clarifying the nature of stochastic fluctuations and accumulation processes in spontaneous movements. *Frontiers in Psychology*, *14*, 1271180. https://doi.org/10.3389/fpsyg.2023.1271180

Boring, E. G. (1950). *A history of experimental psychology* (2nd ed.). Appleton-Century-Crofts.

Breitmeyer, B. G. (1985). Problems with the psychophysics of intention. *Behavioral and Brain Sciences*, *8*(4), 539–540. https://doi.org/10.1017/S0140525X00044915

Dennett, D. C. (1991). *Consciousness explained*. Little, Brown.

Dirnberger, G., Lang, W., & Lindinger, G. (2008). A new method to determine temporal variability in the period of pre-movement electroencephalographic activity. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *70*(3), 165–170. https://doi.org/10.1016/j.ijpsycho.2008.08.006

Haggard, P. (2024). An intellectual history of the “Libet experiment”: embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Dominik, T., Mele, A., Schurger, A., & Maoz, U. (2024). Libet's legacy: A primer to the neuroscience of volition. *Neuroscience & Biobehavioral Reviews*, *157*, 105503. https://doi.org/10.1016/j.neubiorev.2023.105503

Fahrenberg, J. (2019). *Wilhelm Wundt*. https://psycharchives.org/en/item/3af463ec-e142-4a92-ab79-420f97e9ffd2

Frith, C. D., & Done, D. J. (1986). Routes to action in reaction time tasks. *Psychological Research*, *48*(3), 169–177. https://doi.org/10.1007/BF00309165

Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences*, *16*(2), 107–117. https://doi.org/10.1002/1520-6696(198004)16:2

Hosseini, M., Zivony, A., Eimer, M., Wyble, B., & Bowman, H. (2024). Transient Attention Gates Access Consciousness: Coupling N2pc and P3 Latencies Using Dynamic Time Warping. *Journal of Neuroscience*, *44*(26). https://doi.org/10.1523/JNEUROSCI.1798-23.2024

Kant, I. (1790–2005). *Critique of judgement (J.H* (Trans.). B. Bernard & N. books., Trans.). https://doi.org/10.1017/CBO9780511804656

Khalighinejad, N., Schurger, A., Desantis, A., Zmigrod, L., & Haggard, P. (2018). Precursor processes of human self-initiated action. *NeuroImage*, *165*, 35–47. https://doi.org/10.1016/j.neuroimage.2017.09.057

Kornhuber, H. H., & Deecke, L. (1965). Changes in the brain potential in voluntary movements and passive movements in man: Readiness potential and reafferent potentials. *Pflugers Archiv Fur Die Gesamte Physiologie Des Menschen Und Der Tiere*, *284*, 1–17.

Laureys, S., & Schiff, N. D. (2012). Coma and consciousness: Paradigms (re)framed by neuroimaging. *NeuroImage*, *61*(2), 478–491. https://doi.org/10.1016/j.neuroimage.2011.12.041

Subjective referral of the timing for a conscious sensory experience: a functional role for the somatosensory specific projection system in man. (1979). *Brain*, *102*(1), 193–224. https://doi.org/10.1093/brain/102.1.193

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

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*(4), 529–539. https://doi.org/10.1017/S0140525X00044903

Libet, B. (2004). *Mind Time: The Temporal Factor in Consciousness* (Vol. 19). Harvard University Press. https://doi.org/10.1027/0269-8803.19.3.232

Logan, G. D. (2015). The Point of No Return. *Quarterly Journal of Experimental Psychology*, *68*(5), 833–857. https://doi.org/10.1080/17470218.2015.1008020

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Makovski, T., Hommel, B., & Humphreys, G. (2014). Early and late selection: Effects of load, dilution and salience. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00248

Mashour, G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, *105*(5), 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

Matsuhashi, M., & Hallett, M. (2008). The timing of the conscious intention to move. *The European Journal of Neuroscience*, *28*(11), 2344–2351. https://doi.org/10.1111/j.1460-9568.2008.06525.x

Michotte, A., & Prüm, E. (1910). Étude expérimentale sur le choix volontaire et ses antécedents immédiats. *Revue Philosophique de la France Et de l'Etranger*.

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology*, *16*(2), 329–337. https://doi.org/10.1037/h0029815

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/s1364-6613(03)00028-7

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers. *Psychological Review*, *83*(5), 405–408. https://doi.org/10.1037/0033-295X.83.5.405

Parés-Pujolràs, E., Kim, Y.-W., Im, C.-H., & Haggard, P. (2019). Latent awareness: Early conscious access to motor preparation processes is linked to the readiness potential. *NeuroImage*, *202*, 116140. https://doi.org/10.1016/j.neuroimage.2019.116140

Parés-Pujolràs, E., Matić, K., & Haggard, P. (2023). Feeling ready: Neural bases of prospective motor readiness judgements. *Neuroscience of Consciousness*, *2023*(1), 003. https://doi.org/10.1093/nc/niad003

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 in Cognitive Sciences*, *14*(1), 16–21. https://doi.org/10.1016/j.tics.2009.11.001

Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(42), 2904–2913. https://doi.org/10.1073/pnas.1210467109

Seghezzi, S., & Haggard, P. (2023). Volition and "free will." In *The Sage Handbook of Cognitive and Systems Neuroscience*. SAGE Publications Ltd.

Sherrington, C. S. (1906). *The integrative action of the nervous system*. Yale University Press. https://doi.org/10.1037/13798-000

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology*, *117*(11), 2341–2356. https://doi.org/10.1016/j.clinph.2006.04.025

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.

Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*(5), 543–545. https://doi.org/10.1038/nn.2112

Spence, C., & Parise, C. (2010). Prior-entry: A review. *Consciousness and Cognition*, *19*(1), 364–379. https://doi.org/10.1016/j.concog.2009.12.001

Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and attention*. The Macmillan Company.

Tononi, G., Albantakis, L., Boly, M., Cirelli, C., & Koch, C. (2023). *Only what exists can cause: An intrinsic view of free*. https://doi.org/10.48550/arXiv.2206.02069

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. *Social Research*, *44*(4), 757–785.

Vohs, K. D., Schmeichel, B. J., Lohmann, S., Gronau, Q. F., Finley, A. J., Ainsworth, S. E., Alquist, J. L., Baker, M. D., Brizi, A., Bunyi, A., Butschek, G. J., Campbell, C., Capaldi, J., Cau, C., Chambers, H., Chatzisarantis, N. L. D., Christensen, W. J., Clay, S. L., Curtis, J., & Albarracín, D. (2021). A Multisite Preregistered Paradigmatic Test of the Ego-Depletion Effect. *Psychological Science*, *32*(10), 1566–1581. https://doi.org/10.1177/0956797621989733

Ward, J. (Ed.). (1886). Psychology. In *Entry in Encyclopedia Britannica* (Vol. 20). Encyclopaedia Britannica, Incorporated.

Wegner, D. M. (2002). *The illusion of conscious will*. MIT Press. https://doi.org/10.7551/mitpress/3650.001.0001

Wontorra, H., Lorenz, G., & Merkel, J. (2011). *Early Apparatus-based Experimental Psychology, Primarily at Wilhelm Wundt's Leipzig Institute*. https://www.semanticscholar.org/paper/EARLY-APPARATUS-BASED-EXPERIMENTAL-PSYCHOLOGY%252C-AT-Wontorra-Lorenz/6bb8adba42eef1aecc7f3a9815067dd217ce0df1

Haggard, P. (2024). An intellectual history of the "Libet experiment": embedding the neuroscience of free will. In *Proceedings of the Paris Institute for Advanced Study* (Vol. 21). https://doi.org/10.5281/zenodo.13341982
2024/4 - paris-ias-ideas - Article No.13. Freely available at https://paris.pias.science/article/an-intellectual-history-of-the-libet-experiment-embedding-the-neuroscience-of-free-will - ISSN 2826-2832/© 2025 Haggard P.