



The Libet experiment and its implications for conscious will

Peter G.H. Clarke

Summary

A famous experiment of Benjamin Libet and his colleagues has been interpreted as showing that our brains initiate voluntary movements before we are aware of having decided to move, and that this calls into question the efficacy of our wills. These claims have been contested by many neuroscientists and philosophers. This paper provides an introduction to the controversy.

The neurophysiological experiments of Benjamin Libet and his collaborators in the 1980s¹ have been interpreted by the authors and many others as showing that our brains initiate conscious voluntary movements as well as the will to move before we are consciously aware of the will to move. I shall refer to this claim as the *Libet claim* for brevity. It is controversial, but if valid would have important implications for our understanding of how the mind relates to the brain and for the role of conscious will in the performance of voluntary actions. Before going into details about the Libet experiment, I must first provide some information about the mind-brain relationship and the neurophysiology of voluntary movement.

The mind-brain relationship and the Libet claim

It is generally accepted that the electrical activity of our brains underlies our conscious thought, including our decision making. How a physical thing, the brain, can be the basis of consciousness is a subject of debate that has given rise to many different philosophical positions, but these can be grouped in two main categories: dualism and monism.

So great was the influence of Descartes on western philosophy that, from the late seventeenth century until around 1950 or so, most westerners accepted some form of interactive dualism, involving an immaterial soul acting on a material brain. Since then this view has lost favour, for a variety of reasons, including the arguments of philosophers such as Ryle, Place and Feigl. In addition, atheistic materialists rejected it because it invokes a nonmaterial entity, but so did most Christian academics, because advances in the analysis of biblical texts in the mid twentieth century and since tended to support a monistic conception of man, not a dualistic one.² This realisation was not entirely revolutionary, because there had always been a monistic strand in Christian thought due to the influence of Thomas Aquinas. Thus, during his Gifford Lectures in 1956-57, Anglican theologian Austin Farrer criticised the dualistic views of neurobiologist (and future Nobel prize-winner) John Eccles, writing:



About the Author

Dr Peter Clarke was an associate professor at the University of Lausanne, Switzerland, until 2012. He taught anatomy and neuroscience and did research on the mechanisms of neuronal death. His main focus is now on the philosophical implications of neuroscience. He is an associate editor of *Science and Christian Belief* and is a member of The Faraday Institute Advisory Board. He is the author of *Dieu, l'homme et le cerveau* (Croire Pocket, 2012).

We will have nothing to do with the fantastic suggestion, that what the supersensitive 'reactors' in the cortex react to, is the initiative of a virtually disembodied soul. To what, then, are we to say that they do react? What else, than to the motions of the embodied soul, that is to say, other motions in the same nervous system?³

For these reasons, with a few exceptions,⁴ most modern philosophers and neuroscientists, whether theist or atheist, accept some form of monism, but this does not have to involve eliminative materialism that rejects mind as illusory. Many theists, agnostics and atheists adopt more moderate monist positions such as *two-aspect monism*, according to which our subjective, first-personal, account of our inner life and neuroscience's objective, third-personal account of our brain's activity refer to complementary aspects of a single entity.⁵ An alternative view is the *mind-brain identity theory*, according to which the mind and the brain's activity are considered to be the same entity, not two aspects of the same entity. I prefer two-aspect monism because mind-brain identity seems to me linguistically problematic, but the two formulations make identical predictions at the level of brain function.

A striking aspect of the Libet claim is that it goes against the main versions of both dualism and monism. Cartesian dualism predicts that mind events should *precede* brain events, since the nonphysical mind (or soul etc.) is considered to be the real source of our decisions. Two-aspect monism and mind-brain identity theory both predict that mind and brain events should be *synchronous*,

¹ Libet, B., Gleason, C.A., Wright, E.W. & Pearl, D.K. 'Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act', *Brain* (1983)106: 623-642.

² Green, J.B. *Body, Soul and Human Life: The Nature of Humanity in the Bible*, Carlisle: Paternoster (2008).

³ Farrer, A. *The Freedom of the Will*, London: A & C Black (1958), p. 87.

⁴ Goetz, S. and Taliaferro, C. *A Brief History of the Soul*, Chichester, UK: Wiley-Blackwell (2011).

⁵ Nagel, T. *The View From Nowhere*, Oxford: Oxford University Press (1986), chap. 3, p. 28; Jeeves, M. & Brown, W.S. *Neuroscience, Psychology and Religion*, West Conshohocken, PA: Templeton Foundation Press (2009).

since mind-level descriptions and brain-level descriptions are considered complementary (and equally valid) accounts of the same processes. But if brain events come first, this would support *epiphenomenalism*, the view that mind events are mere by-products of brain events, with no causal role. This would deny the causal efficacy of conscious will.

The neurophysiology of voluntary movement

It is important to be clear about what is, and is not, being claimed when a movement is called voluntary. Even though these movements involve, by definition, an act of conscious will, that is not to say that every aspect of the movement is conscious or willed. For example, the movements of a tennis player as she serves are voluntary, but their control involves many automatic subroutines in the cerebellum and elsewhere. Furthermore, to claim that conscious acts of will initiate voluntary movements is not to deny that the acts of will arise out of brain processes that are largely unconscious.⁶

What is the nature of the 'I' (or self) that willed the movement and performed it? The use of such terms does not imply dualism. The 'I' (or self, or mind etc.) is generally conceived as being embodied in (or emerging from) the brain's activity.

The neural circuits involved in voluntary motor control are exceedingly complicated, and I here give only some simplified information that is necessary for understanding the Libet experiment. Voluntary movements are controlled primarily by the motor cortex (in the back part of the frontal lobe – Fig. 1) but in cooperation with many other motor centres including the basal ganglia and the cerebellum. Motor commands are sent from the primary motor cortex (and to some extent from other areas) to motoneurons in the brainstem and spinal cord, which in turn control the muscles. The initiation and programming of movements depend on activity in many areas including the supplementary motor area (I) and the preSMA, and several areas in the parietal cortex. These areas feed directly or indirectly into the premotor cortex and motor cortex. Electrical stimulation of the motor areas produces movements, but not the will to move. In contrast, electrical stimulation of areas BA-39 and BA-40 in the parietal lobe (Fig. 1) elicits the will to move, but does not cause a movement.⁷

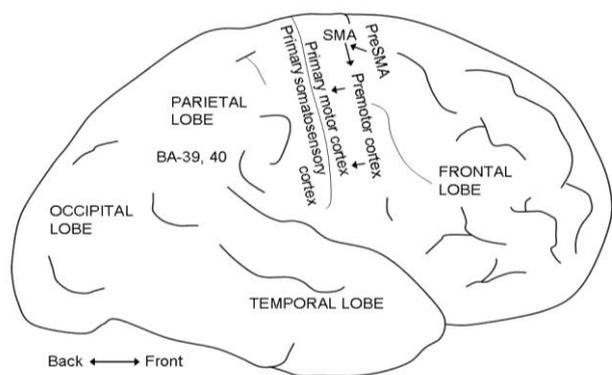


Fig. 1. The brain's cerebral cortex, viewed from the right side. SMA: supplementary motor area. BA: Brodmann's area.

The Libet experiment, a challenge to the role of conscious will

An important background to the Libet experiment was the discovery in the 1960s that, before people make a voluntary movement, there is a slow build-up of electrical potential measured from the skull over the motor cortex, beginning as

much as a second earlier for simple movements and even longer for complex series of movements.⁸ This electrical change is called the readiness potential (RP).

Libet was interested in the relative timing of the RP compared with the movement and the conscious decision to move. He therefore asked his experimental subjects to perform simple movements, in most cases flexion of the fingers or wrist, and to estimate the time of conscious awareness of the urge (or will or decision) to move (W) by reporting the position of a spot moving in a circle on an oscilloscope screen. They were told to perform the movement whenever they felt like doing so, and to pay close attention to the time when they were first aware of the 'urge to move'. He also recorded the RP by electroencephalography, and the time of the movement itself was estimated from the electromyogram. Libet found that time W came only about 200 msec before the movement, whereas the RP began much earlier, usually about 550 msec before the movement (Fig. 2). The fact that the change in brain potential occurred before the conscious decision was interpreted by Libet and by many commentators to imply that our conscious decision to act is not the true cause of the movement. They deduced that conscious will is too slow to make things happen, and that volitional acts must result from unconscious processes in the brain, not from conscious willing. This seemed to imply that our intuitive notion of conscious will must be an illusion.

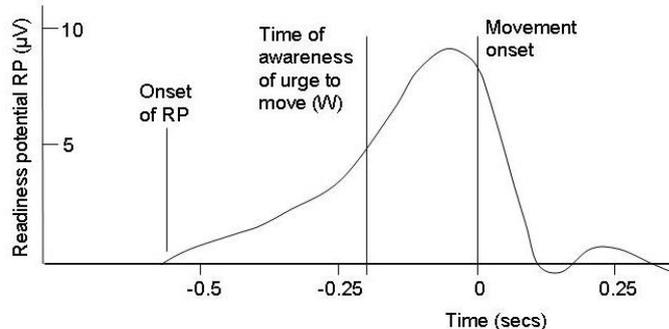


Fig. 2. Schematised readiness potentials (RPs) preceding self-initiated voluntary acts, as in the Libet experiment. Since these scalp-recorded potentials are of only about 10 μ V, smaller than the background EEG, the experimenters had to average about 40 raw recordings to obtain reproducible results. We follow Libet et al. in using the term 'Time of awareness of urge to move' and in designating it by W (for will).

There appeared to be a small loophole in that Libet's subjects still had the power to *veto* a movement in the 200 msec between time W and the movement. He therefore argued that even though the *initiation* of the movement was not the result of conscious will, its *vetoing* was. This argument has not attracted great interest, but was supported by eminent free-will philosopher Robert Kane.⁹

The Libet experiment provoked considerable interest and intense controversy, and stimulated further experimentation.

Single neuron recordings during the Libet experiment

It is rarely possible to record from single neurons in the brains of humans, but this can occasionally be done in epilepsy patients using electrodes that have been implanted to localise the zones that cause seizures. Thus, remarkably, Itzhak Fried and his collaborators managed to record from more than 1,000 neurons in the medial

⁶ Gomes, G. 'The timing of conscious experience: a critical review and reinterpretation of Libet's research', *Consciousness & Cognition* (1998) 7: 559-595.

⁷ Desmurget, M., Reilly, K.T., et al. 'Movement intention after parietal cortex stimulation in humans', *Science* (2009) 324, 811-813.

⁸ Kornhuber, H.H. and Deecke, L. 'Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale', *Pflügers Archiv*(1965) 284: 1-17.

⁹ Kane, R. *The Significance of Free Will*, New York / Oxford: Oxford University Press (1996), p. 232.

frontal cortex of epilepsy patients (and especially in the supplementary motor area, which generates most of the early part of the RP) as they performed the Libet experiment. It was found that a few neurons changed their firing rate (by an increase or a decrease) almost 1.5 secs before time W, and more and more neurons did so over the following 1.5 secs, with about 25% of the neurons firing several tenths of a second before W. The authors conclude that their findings support the view that the experience of will emerges as the *culmination* of premotor activity starting several hundreds of msec before awareness.¹⁰

Criticisms of the Libet claim

Despite the fame of the Libet experiment and its frequent acceptance in popular and semi-popular writings, it has been the subject of intense controversy. Indeed, most specialists in the philosophy of free will who have addressed the Libet claim have rejected it.¹¹ Most of the criticisms focused on difficulties of judging the time of awareness, of interpreting the RP, or of philosophical interpretation, as is discussed below.

Problems of judging the time of awareness

It was central to Libet's claim that the readiness potential began distinctly before time W. The published data of several groups do indeed support this claim, but critics have objected to the use of subjective recall after the event, because there is evidence that this can be very unreliable. Furthermore, those such as Alfred Mele¹² who have tried the experiment for themselves have found that W is difficult to define. I have done this too, and you may wish to try it using a 'clock' available on the web.¹³ When I try this, I find it very hard to judge the precise time when I decided to move my finger/wrist. It would be useful to quantify the reliability of our judgements, but this is difficult for a purely subjective decision. For this reason, several research groups have instead measured the reliability of timing judgements for *perceptual* events, which is easier to do. Results have been variable, but several groups found serious biases,¹⁴ raising doubts about the interpretation of the Libet experiment. A different critique of the timing was made by Dennett and Kinsbourne,¹⁵ who point out that Libet's experiment involves an attention shift from the participants' subjective intention to the clock, which may have introduced temporal mismatches between the felt experience of will and the perceived position of the clock hand.

To try to solve these problems, Matsuhashi and Hallett devised an alternative methodology for estimating time W. They found that the RP (which they called BP1) occurred before W in only about two thirds of the subjects; worse, the lateralised RP (LRP) that we shall discuss below, always occurred *after* W.¹⁶

In view of the controversy about the measurement of subjective timing, considerable attention was devoted in the public media to a paper published in Nature Neuroscience that used brain scanning technology (functional magnetic resonance imaging – fMRI) in a Libet-like experimental paradigm, and included in the summary a claim that a 'decision can be encoded in brain activity of prefrontal and parietal cortex almost 10 sec

before it enters awareness'.¹⁷ After all the subtle debate about a few hundreds of milliseconds, 10 sec was an enormous amount of time, and the wording of the abstract gave the impression that the temporal priority of the neural decision with respect to the subjective one was finally established. I assume that some journalists and bloggers only had access to the abstract (available free on the web) and not to the full paper, because the main text made only the much weaker claim that the activity of prefrontal and parietal cortex was *correlated with the decision* (to use the left or right hand) with 60% prediction accuracy, up to 10 sec before the conscious decision. That is very different! To reflect a neural decision, the correlation would need to be at 100%, not 60%. The paper provided valuable information about brain activity leading ultimately to a decision, but did nothing to rescue the Libet experiment from the criticisms about timing.

The overall conclusion on timing has to be that the problems have not so far been resolved.

Doubts as to whether the readiness potential reflects a decision to move

The Libet claim assumes that the RP reflects a neural 'decision' to move, and that the neural activity underlying the RP causes both the will to move and the movement. Even if such causality could be demonstrated, this would not strictly be sufficient to validate the Libet claim, because the decision must presumably be caused by a chain of preceding neural events, and the RP might reflect some of these. But the Libet claim certainly assumes causality. This is part of the claim, and it has never been proved.

To be precise, we are really talking about the *earliest part* of the RP, because the timing argument focuses on the RP's onset. To attribute such a decisional and causal role to this *earliest part* of the RP seems surprising, because it originates mainly in the SMA (Fig. 1), which has been known for more than thirty years to be strongly activated when subjects 'programme' (imagine) a complex movement *without actually performing it*.¹⁸ This is not to deny that activity in SMA can cause movements in some cases, such as when it is stimulated electrically, but it cannot be assumed that the earliest part of the RP necessarily reflects neural processes underlying a decision to move. And there are at least six specific reasons to doubt this.

First, even though electrical stimulation of the SMA can cause movements, it does not cause a will to move, which requires stimulation of parietal areas.¹⁹ This suggests that the RP does not cause the will to move.

Second, if the RP truly caused the conscious will and the movement, one would expect trial-to-trial variations in the onset of the RP to correlate with trial-to-trial variations in time W; that is to say that trials with an early RP should also have an early W. Haggard and Eimer tested this, using a variant of the Libet experiment, and found there was little correlation, ruling out the RP as a cause of the will or decision to move. They did, however, find that the 'lateralized readiness potential' (LRP: i.e. the RP from the cortex on the opposite side relative to the movement minus the RP from the same side) gave a positive correlation, suggesting that the brain processes underlying the LRP might cause the will to move.²⁰ At the time, their paper did not seem to challenge the Libet claim, because the LRP seemed to fulfil the role formerly attributed to the RP. However, the LRP occurs later than the RP, and subsequent experiments have sometimes found that the LRP occurs even after

¹⁰ Fried, I., Mukamel, R. & Kreiman, G. 'Internally generated preactivation of single neurons in human medial frontal cortex predicts volition', *Neuron* (2011) 69: 548-562.

¹¹ Bayne, T. 'Libet and the case for free will scepticism', in Swinburne, R. (ed.) *Free Will and Modern Science*, Oxford: Oxford University Press (2011).

¹² Mele, A. R. *Effective Intentions: The Power of Conscious Will*, New York / Oxford: Oxford University Press (2009).

¹³ There is a suitable clock at

http://www.informationphilosopher.com/freedom/libet_experiments.html

¹⁴ Danquah, A.N., Farrell, M.J. & O'Boyle, D.J. 'Biases in the subjective timing of perceptual events: Libet et al. (1983) revisited', *Consciousness & Cognition* (2008) 17: 616-627.

¹⁵ Dennett, D.C. & Kinsbourne, M. 'Time and the observer', *Behavioral and Brain Sciences* (1992) 15: 183-247.

¹⁶ Matsuhashi M. & Hallett, M. 'The timing of the conscious intention to move', *European Journal of Neuroscience* (2008) 28: 2344-2351.

¹⁷ Soon, C.S., Brass, M., Heinze, H.J. & Haynes, J.D. 'Unconscious determinants of free decisions in the human brain', *Nature Neuroscience* (2008) 11: 543-545.

¹⁸ Roland, P.E., Larsen, B., Lassen, N.A. & Skinhoj, E. 'Supplementary motor area and other cortical areas in organization of voluntary movements in man', *Journal of Neurophysiology* (1980) 43: 118-136.

¹⁹ Desmurget et al. *op. cit.*, (7).

²⁰ Haggard P. & Eimer M. 'On the relation between brain potentials and the awareness of voluntary movements', *Experimental Brain Research* (1999) 126: 128-133.

time W as is discussed above,²¹ so the LRP seems a fragile candidate to replace the RP.

Third, Alfred Mele has pointed out a flaw in Libet's experimental paradigm that vitiates attempts to deduce a causal influence between the RP and the movement (and the will to move).²² In all Libet's experiments, the permanent storage of electroencephalographic data was triggered by the finger/wrist movements. This was necessary as part of the averaging procedure that is necessary to detect the RP, which would otherwise be masked by other concurrent activity in the EEG. If there was no movement, the data were not stored, so any RPs that occurred without being followed by movements would not have been detected. If such RPs without movement did occur, then RPs are not sufficient to cause movements, and more probably reflected brain activity occurring prior to the decision to move. This possibility is difficult to evaluate, because the averaging procedure has to be triggered at a moment defined by the movement.

Fourth, experiments by Hermann et al. cast further doubt on the interpretation of the RP as causally related to the decision and movement.²³ These researchers used a modified version of the Libet experimental paradigm, in which the participants were instructed to press one of two buttons, depending on a presented stimulus. An RP occurred well before the motor response, as in the Libet experiment. But, importantly, it occurred even before the stimulus presentation, so it clearly did not reflect a decision as to which button to press. The authors argue that the RP does not specifically determine the movement, but may reflect a general expectation (which is indeed what the RP was initially thought by Kornhuber and Deecke to reflect, not a decision but a state of readiness, hence its name).

Fifth, Trevena and Miller devised a modified version of the Libet experiment in which participants made spontaneous decisions to move, or not, and found that the RP was no stronger before a decision to move than before a decision not to move, which is not what one would expect if the RP reflected a neural decision to move.²⁴

Sixth, computational analysis suggests that the neural decision to move occurs only very late during the time-course of the RP, not at its onset.²⁵

Debate about the philosophical interpretation

Even if the Libet claim is accepted – which is very controversial, as we have seen – there is also debate about the philosophical interpretation.

I have here used systematically the term 'conscious will' rather than 'free will' to avoid the broader philosophical associations of the latter term. Nevertheless, many supporters of the Libet claim, including Libet himself²⁶, have used the term 'free will'. This has aroused further controversy, because many critics have argued that Libet's experimental paradigm was

irrelevant to the question of free will. When we talk about free will, we are usually referring to choices among a variety of options, often with moral implications, and this may require careful deliberation over a period of minutes or hours or days. The Libet experiment is just the opposite. The subject was not making a moral decision, and was not even deciding *whether* to move, but only *when*. Moreover, the subjects were specifically instructed not to deliberate but to act spontaneously, and in their original 1983 paper Libet et al. explicitly pointed out that their conclusions applied only to spontaneous, rapidly performed movements.²⁷ Thus, even if we accept the debatable claim that the finger/wrist movements in the Libet experiment were not the result of conscious will, this conclusion cannot automatically be extended to situations for which the term free will would normally be applied.

Another problem is that those who support an anti-free-will interpretation appear to have in mind only rather marginal notions of free will. For example, in a review on the neuroscience of volition, neurobiologist Haggard, a former collaborator of Libet and leading protagonist of the anti-free-will interpretation, mentions the possibility that the brain's circuits might be influenced by 'an unspecified and uncaused cause (the "will")'. Haggard rejects this view, and concludes the article by stating that 'modern neuroscience is shifting towards a view of voluntary action being based on specific brain processes...'.²⁸ This gives the impression that 'modern neuroscience' is gradually triumphing against the illusion of free will, but this is confusing for at least two reasons. First, only a tiny minority of modern philosophers conceive of the will as an 'uncaused cause', so why use such a marginal definition? Second, the words about modern neuroscience's 'shifting towards a view of voluntary action being based on specific brain processes' are strange, because this has been the standard view in neuroscience for over half a century. In the same review, Haggard states that the Libet experiment 'seems to disprove the everyday concept of "free will"'; his reference to 'everyday concept' suggests that he recognises that this challenge does not extend to more sophisticated concepts of free will.

Conclusion

Libet's 1983 experiment reported that brain activity (the RP) reflecting a decision to flex a finger or wrist occurred several hundred milliseconds before the subject became aware of her decision (or urge or will) to move. This has been interpreted, controversially, to suggest that our subjective impression that our conscious wills initiate the movement is illusory. Libet accepted this interpretation, but maintained that conscious will can still play a genuine role in the vetoing of initiated acts.

Many neuroscientists and most philosophers contest the claims about the supposed inefficacy of conscious will, and this paper summarises their arguments. At the neurophysiological level, it has not been shown convincingly that a neural 'decision' sufficient to cause the movement occurs before the time of awareness of the decision to move. Even if this could be shown, it would not undermine the conceptions of free will that are defended by most philosophers.

Acknowledgments

The author is grateful to Martyn Frame and Stuart Judge for their helpful comments on an earlier draft of this paper.

²¹ Matsushashi & Hallett *op. cit.*, (16).

²² Mele *op. cit.*, (12).

²³ Herrmann, C.S., Pauen, M., Min, B.K., Busch, N.A. & Rieger, J.W. 'Analysis of a choice-reaction task yields a new interpretation of Libet's experiments', *Intl J. Psychophysiol.* (2008) 67, 151-157.

²⁴ Trevena, J. & Miller, J. 'Brain preparation before a voluntary action: evidence against unconscious movement initiation', *Consciousness & Cognition* (2010) 19, 447-456.

²⁵ Schurger, A., Sitt, J.D. & Dehaene, S. 'An accumulator model for spontaneous neural activity prior to self-initiated movement', *Proc. Natl. Acad. Sci. USA* doi:10.1073/pnas.1210467109 (2012).

²⁶ Libet, B. *Mind Time*, Cambridge Mass / London, UK: Harvard University Press (2004).

²⁷ Libet et al. *op. cit.*, (1).

²⁸ Haggard, P. 'Human volition: towards a neuroscience of will', *Nature Reviews Neuroscience* (2008) 9: 934-946.

The Faraday Papers

The Faraday Papers are published by the Faraday Institute for Science and Religion, St Edmund's College, Cambridge, CB3 0BN, UK, a charitable organisation for education and research (www.faraday-institute.org). The opinions expressed are those of authors and do not necessarily represent the views of the Institute. The Faraday Papers address a broad range of topics related to the interactions between science and religion. A full list of current Faraday Papers can be viewed at www.faraday-institute.org from where free copies can be down-loaded in pdf format. Print copies like this one can also be obtained in bulk quantities of ten or more at £1.50 per 10 copies + postage. Secure on-line ordering details are at www.faraday-institute.org.

Publication date: February 2013. © The Faraday Institute for Science and Religion