

Does neuroscience prove that free will does not exist?

Daan Evers

(Lecture for students of the Legal Research Master

Utrecht University, 26/01/2012)

1. Responsibility and justice

Dutch criminal and civil law agree that one cannot be convicted of a crime unless one is *responsible* for it. And it's natural to think that one cannot be responsible for an act unless one freely decided to do it. So if free will does not exist, this might have drastic consequences for our legal system. Since many legal codes are based on similar assumptions, the debate about free will appears to have enormous societal relevance worldwide.

Still, it's not *obvious* that the discovery that free will does not exist would have drastic consequences for the legal system. Dutch criminal law provides a set of conditions which defeat a person's responsibility, such as *force majeure*, youthfulness and mental illness. If free will does not exist, this needn't mean that such requirements are always met. For example, not everyone would be mentally ill if free will did not exist. It's not even clear that the criterion of *force majeure* would be met, as this is ordinarily understood in terms of unforeseen or uncontrollable events external to the agent (like hurricanes or strikes).

Of course, the law might also involve a *general* presumption of responsibility, such that suspects are appropriate objects of moral blame or praise in a sense *not* exclusively defined by criteria like the ones just mentioned. This would make the debate about free will more clearly relevant. For whatever free will may be exactly, it is often taken to be the ability to control one's actions in such a way as to be morally responsible for them.

So it might be of more than purely philosophical interest that some neuroscientists claim to have discovered that free will does not exist, or is unlikely to exist. This is of course not an original opinion, as many philosophers have denied its reality before, but what *is* new is the application of empirical methods to an ancient philosophical question. In this talk I will discuss some neuroscientific results which have been taken to undermine free will. I will argue that these do not provide strong evidence against the existence of free will. If so, neuroscience does not (yet) give us any special grounds to be concerned about the legal system.

2. Benjamin Libet

It's probably fair to say that the neuroscientific interest in free will began with the now deceased Benjamin Libet. Libet (1983, 1985) performed experiments in which he asked subjects to perform a simple motor action, such as flexing one's wrist or pressing a button, within a certain period of time. Subjects were asked to do so spontaneously, whenever they felt like doing it. While performing this task, they were asked to watch a fast-moving dial on a clock face. Subjects were asked to

remember the position of the dial when they first became aware of what Libet described as the *decision, urge or intention* to act.

So subjects had to pay attention to two things: first, the occurrence of their own decision and, second, the position of the dial on the clock when they decided. What Libet did while subjects were performing this task was to measure electrical activity in regions of the brain associated with motoric action. And what he found was the following: a delay between the onset of increased electrical activity in the motor cortex and subjects' reports of being (first) aware of their decision (urge, intention) to act. On average, the electrical activity starting ramping up 550 milliseconds before execution of the act, while the conscious awareness of the decision occurred around 200 milliseconds before execution. So there was an increase in electrical activity in the motor cortex which preceded the subjects' consciousness of their decision by 350 milliseconds. This result is often described in terms of the brain 'deciding' to perform a certain action prior to the subject's awareness.

Two kinds of conclusions have been drawn from experiments like these: the first is that the conscious decision to move your arm is itself caused by unconscious brain processes, in particular the measured activity in the motor cortex which is known as *readiness potential*. The second conclusion that has been drawn is that conscious decisions are causally irrelevant to action, that they have no causal role at all. For example, Roediger et al. (2008) said that Libet's findings contradict the 'naive view' that

'conscious intention causes action. Clearly conscious intention cannot cause an action if a neural event that precedes and correlates with the action comes before conscious intention' (2008, p. 208).

But this conclusion certainly does not follow from Libet's results. Even if some event E2 is caused by another event E1, this doesn't mean that E2 plays no role in causing subsequent events. If the brick's hitting the window was itself caused by my throwing it, the brick's hitting the window can still cause the glass to break. So these results don't help to show that conscious decisions are, as philosophers say, epiphenomenal: causally irrelevant by-products of processes that do the causal work.

The first conclusion is, on the face of it, more credible: that conscious decisions are themselves caused by prior unconscious processes. But why should this conclusion, if true, undermine free will? People who think that free will is threatened clearly think that a freely willed action has to *originate* in the subject's conscious decision. And here's one reason why you might think this is required: a freely willed action is an action over which the subject has control. But if the subject's action is ultimately caused by unconscious processes, then s/he lacks control over what s/he does (after all, she is not even aware of what is going on in her brain). So in order to exercise control over your actions, they have to originate in uncaused conscious states, like decisions.

I'm inclined to think this is correct, with certain caveats which we needn't be concerned with now.¹ So do Libet's original experiments warrant the conclusion that the decision to flex one's wrist or push the button was caused by prior unconscious neural activity (the readiness potential)? No, for at least three reasons. First, because of the way the experiment was set up, there are only recordings of the neural activity preceding *executions* of the action. So nobody knows whether the same readiness potential can occur without subsequent decision to act (Mele 2009). Second, as Tim Bayne points out, 'we can make willed responses to stimuli in very much less than 550 milliseconds' (2011, p. 40). This suggests that it might well be possible for the readiness potential to be interrupted by events at some later stage, including conscious decisions to act otherwise. Third, the readiness potential measured by Libet is *generalized*, meaning that it is symmetrically distributed over the brain's two hemispheres. That generalized readiness potential has been found prior to many different kinds of motoric actions and decisions (Haggard & Eimer 1999). So it is not specific enough to determine what the agent will decide. This suggests that something else determines it, which might be the agent's spontaneous activity.

So it would certainly be hasty to conclude from Libet's results that free will does not exist. This would be hasty even if you accept that it's a requirement on freely willed action that it originates in uncaused conscious decisions. For Libet did

¹ The main one being that not *everything* I do has to be caused by (uncaused) conscious decisions. When I walk, I don't consciously decide to put my right leg forward, then my left, etc. It is an automatic process. This does not make my act of walking unfree, provided I can choose to start walking and decide to stop as well.

not even show that the simple (unreasoned) decisions in his experiment were caused by unconscious processes, let alone all decisions in real life.

It should be noted that Libet himself did not believe that his experiments undermined free will, even though he did accept that the decision to flex a wrist or push a button was caused by the readiness potential. He did not believe they undermined free will because he thought it possible for the subject to veto the decision in a 150 ms-long period between its occurrence and execution of the act (Libet also tried to prove this experimentally). So Libet would resist the argument from control which I sketched earlier, by pointing out that the agent can still veto decisions that are caused by unconscious neural events.

3. Patrick Haggard and Martin Eimer

I said that the generalized readiness potential measured in Libet's experiments was too coarse-grained to be a plausible cause of particular decisions. After all, it occurs prior to any motoric action whatsoever, so many different decisions are compatible with it. This was pointed out in an article from 1999 by neuroscientists Patrick Haggard and Martin Eimer.

Haggard and Eimer gave an additional reason to doubt that Libet's readiness potential was the cause of the decision to move. This reason was that the timing of the increase of electrical activity did not seem to covary with the timing of the decision of the subject. So earlier increases did not necessarily correspond to earlier decisions by the subjects. Likewise, later rampings-up did not necessarily correspond to later decisions. Haggard and Eimer took this to be a reason to doubt

that the generalized readiness potential was the cause of the decision. In order to back this up, they appealed to a rule of John Stuart Mill's for identifying causes. Mill roughly said that you can infer an effect to be caused by a certain phenomenon just in case changes in the effect covary with changes in that phenomenon (provided the effect and the phenomenon have no common cause). Libet's readiness potential did not seem to satisfy this criterion.

But Haggard and Eimer argued that the timing of a *later* stage of the readiness potential, the so-called lateralized readiness potential, *did* covary with the timing of decisions and therefore might be their (unconscious) cause. This lateralized readiness potential, which occurs only in one of the brain's hemispheres, precedes the report of the decision by much less than the onset of the generalized readiness potential, although it doesn't quite coincide with it either.

They base this observation on eight subjects who were asked to choose to press either a left button (with their left hand) or a right button (with their right hand). Otherwise, the setup was similar to Libet's. Haggard and Eimer compared the timing of the lateralized readiness potential with the timing of the subjects' conscious decisions (or more precisely, their reports of such decisions). What they found was that when you average out over a number of trials, earlier lateralized readiness potentials were followed by earlier decisions, and later lateralized readiness potentials were followed by later decisions.

The results are shown in the following table (from Haggard & Eimer 1999, p. 132):

Subject	LRP onset		W judgment	
	Early W trials (ms)	Late W trials (ms)	Early W trials (ms)	Late W trials (ms)
1	-870	-805	-542	-351
2	-800	-900	-462	-257
3	-1320	-905	-494	-243
4	-1100	-1100	-273	-118
5	-800	-810	-940	-4
6	-800	-185	-231	-80
7	-1030	-600	-370	-172
8	-530	-400	-984	-253

On the left you can see that there were eight participants in the experiment. The columns underneath 'LRP onset' represent (for each subject) the average time that the lateralized readiness potential started ramping up before execution of the act. The left column represents the average time of the onset in the case of early decision reports and the right column represents the average time of the onset in the case of late decision reports. The columns underneath 'W judgment' (= decision report) represent (for each subject) the average time that the subject reported awareness of the conscious decision to press a button before execution of the act. Again, the left column represents early decision reports, while the right column represents late decision reports.

Haggard and Eimer claim that these results are consistent with Mill's principle that you can infer something to be the effect of a certain phenomenon just in case the two covary. That is to say: if the one is earlier, the other should be earlier too, and if the one is later, then the other should be later too. But this does not appear to be the case.

Consider subject number 5. Number 5's LRP onset in early decision reports occurs (on average) 800 milliseconds before she presses the button. But her decision report occurs (on average) 940 milliseconds before pressing the button. So the report *precedes* the onset of the LRP. But if the LRP were the cause of the decision, then the decision report should occur *after* the LRP.

Now consider subject number 8. Number 8's LRP onset in early decision reports occurs (on average) 530 milliseconds before pressing the button. But her decision report occurs (on average) 984 milliseconds before pressing the button. So, again, the report of the conscious decision *precedes* the onset of the LRP, which is supposed to be the cause of the decision. Surely causes have to come before their effects, not the other way around. So in two out of eight subjects, certain averages are not consistent with Mill's principle. So by the authors' own lights, the lateralized readiness potential is not plausibly the cause of the decision to press the button.

Also notice great differences in the interval between the onset of the LRP and the decision reports. Take subject number 4. In the late decision trial, (second column), the LRP onset occurs (on average) 1100 milliseconds before execution of the action. Her decision report occurs 118 milliseconds before. So there is an interval of 982 milliseconds between the onset of the LRP and the decision report. In the case of subject number 6, the interval is (on average) only 105 milliseconds. So this is another respect in which the onset of the LRP and the decision report do not covary (in fact, the intervals are different for all subjects). So, again, by the authors' own light, they cannot reasonably infer that the LRP is the cause of the decision to press the button.

Finally, the numbers in the table represent averages over a number of trials. This means that even if earlier LRP onsets correspond with earlier decision reports on average (and later ones with later ones), there may be individual trials for each subject in which this is not true. If that is indeed the case, then there are even more violations of Mill's principle than we can gather from the table.

Haggard and Eimer do not claim that lateralized readiness potentials *are* the causes of conscious decisions in their experiments, but they do think this is a possibility. However, if you go by the principle that an effect can be caused by some phenomenon only if that phenomenon covaries with the effect, this does not seem to be a possibility.

There is another reason to be cautious about taking the lateralized readiness potential as a cause of the conscious decision to press left or right. With respect to Libet's own experiments, Dennett and Marcel (1992) noted that reports of conscious decisions might be significantly later than they actually took place, since subjects had to divide their attention between the clock and their own inner mental workings. The idea is that processing *that* one has decided takes a certain time, as does processing the visual information from the clock. This might lead to mismatches between the reports and the actual timing of events. This problem is unlikely to throw into doubt the precedence in time of the *generalized* readiness potential (Bayne 2011), but since the gap between the *lateralized* readiness potential and the report of the decision is much smaller, it may haunt an interpretation of the LRP as the cause of the conscious decision.

I conclude that neither Libet nor Haggard and Eimer provide strong reason to worry about the existence of free will.

4. John-Dylan Haynes

Neuroscientists often complain that philosophers keep flogging Benjamin Libet, whose work is from the eighties, while there are much more recent findings which would provide strong evidence against the existence of free will. My colleague Niels van Miltenburg and I were accused of this ourselves by a psychologist at Utrecht University, when we published an article in the Dutch newspaper *De Volkskrant*. (In a way, I welcome this complaint, insofar as it reflects agreement that Libet did not provide strong evidence against free will.) To support his point, the psychologist sent us a recent article by Soon et al., among whose authors is the well-known neuroscientist John-Dylan Haynes.

The German newspaper *Die Zeit* published an article on Haynes in which we find the following:

'manche Philosophen versuchen, die Idee des frei schaltenden und waltenden Geistes zu retten", sagt Haynes. "Aber am Determinismus, dem zwingenden Nacheinander von Ursache und Wirkung, führt kein Weg vorbei." Die Philosophie solle sich eher mit den Konsequenzen aus den Befunden der Hirnforschung beschäftigen, als verlorenes Terrain weiter zu verteidigen, schlägt er vor. Was folgt für unser Menschsein aus der Tatsache, dass Psyche und Gehirn eine untrennbare Einheit bilden? Wie lebt es sich ohne freien Willen?' (*Die Zeit*, October 2010)

So, rather than trying to defend antiquated ideas, philosophers should think about the consequences of the discovery that free will does not exist. One would expect someone who says this to have solid evidence that free will does not exist. So what is this evidence?

Haynes and co-authors (Soon et al. 2008) did an experiment in which subjects were asked to spontaneously decide to either press a right button (with their right hand) or a left button (with their left hand) whilst lying in an fMRI scanner. Haynes claimed they could predict which button subjects were going to press on the basis of unconscious brain activity occurring up to seven seconds prior to the subject's reported conscious decision of which button to press.

Haynes and his co-authors called these processes 'unconscious determinants of decisions', clearly implying that they *caused* the subsequent decision (this is also suggested by Haynes' talk of the '*zwingenden Nacheinander von Ursache und Wirkung*' in *Die Zeit*. Haynes clearly believes that the causation of decisions by unconscious processes is what undermines free will. Elsewhere he says: 'How can I call a will 'mine' if I don't even know when it occurred and what it has decided to do?' (Smith 2011).

Given that philosophers get vilified for discussing Libet's work instead of more recent findings, it would be ironic if exactly the same arguments apply to the use of these more recent findings in the context of free will. But it seems to me that this is indeed the case.

A basic worry about the idea that Libet's readiness potential was the cause of the decision to act was this: nobody knew if the same readiness potential could occur without subsequent decision. In the case of Haynes, the situation is worse:

we *do* know that the same activity can occur without that decision occurring. This is because Haynes' predictions are only correct 60% of the time. So in 40% of all cases, the same process which supposedly causes the decision occurs without the decision occurring. This means that something else can influence the outcome. And nothing so far rules out that this is the agent's ability to originate a different choice.

Even stranger is the fact that Haynes and co-authors say the following in a study in which they replicate their earlier results:

'Please [...] note that our study cannot provide evidence for a causal relationship between the activation in frontopolar cortex and the decision, e.g. because fMRI measures neural decision-related processes only indirectly and prediction is far from perfect.' (Bode et. al. 2011, p. 9)

So they admit that the study does not provide evidence for a causal relationship between the unconscious brain processes preceding the decision, and the decision. But the idea that this relation is causal motivates Haynes's denial of free will. So if he has no evidence of it, why is he so confident that free will does not exist?

Perhaps Haynes believes that the results *hint* at a causal link, even if they are not hard evidence of it. This may not be unreasonable. After all, there *is* something to explain: if the unconscious processes did not even *influence* the decision, one shouldn't expect Haynes to be able to predict any better than chance. But his results are clearly significant: if he had guessed, he would have got it right 50% of the time (since the choice is between left and right). But, looking at the

unconscious processes prior to choice allowed him to get it right 60% of the time. Is this not a reason to think that the decision is (at least sometimes) caused by the unconscious process?

Well, it certainly seems to be influenced by something of which the subject is not aware. But this may be an artifact of the experimental setup, in the following sense: subjects are asked to *randomly* decide between left or right presses. Since the subject has no reason to choose one way or the other, she can go on autopilot. In that case, she may allow unconscious biases to determine the outcome. If this is right, then the experiment provides little reason to think that decisions made *for reasons* (rather than randomly) are caused by unconscious brain processes. So it's not clear that these results have any bearing on decisions made after conscious deliberation of the pros and cons.

Still, the experiments may have bearing on decisions made for no reason at all. But even here we should be wary of generalization. After all, it's not clear that the kind of task performed in the experiment bears relevant similarities to other kinds of spontaneous decision. So, again, it would be hasty to conclude that free will does not exist.

5. Daniel Wegner

Finally, I'll briefly discuss some ideas of Daniel Wegner's. Wegner wrote a book provocatively titled *The Illusion of Conscious Will*. In it, he defends the idea that although we feel as though we cause our acts by willing them, this is in fact an

illusion. The experience of conscious will is a by-product of neurological events that do all the causal work.

How does Wegner reach his conclusion? Part of his argument refers to Benjamin Libet's experiments. I've already pointed out that Libet's research provides no grounds for claiming that conscious states are causally irrelevant to action. It could at most establish that conscious states are themselves caused by prior neurological events. But this would not show their causal redundancy. Of course, I've also argued that Libet's results do not warrant the conclusion that decisions *are* caused by preceding neurological events. So this won't help Wegner to establish his claim that conscious will is an illusion.

Wegner is also fond of pointing out potential mismatches between our experience of agency and what is actually going on. For example, in the case of ouija boards, people believe they are not pushing the glass (or whatever it is), whereas they really are. Likewise, you can induce the belief that you are consciously controlling something, whereas in fact you're not. You might be familiar with this experience from videogames, when you wrongly believe you are controlling the racing car at the top of the screen, whereas you are in fact controlling the one at the bottom. Wegner takes such phenomena as evidence that the experience of conscious will is realized by different regions of the brain than motoric action. And he may well be right. But it does not seem to follow that the experience of conscious will is always illusory:

'[T]o say that the [brain] systems [underlying consciousness and movement] are not identical is not to say they are unconnected' (Nahmias 2002, p.

532).

But *this* is what Wegner has to show.

Nahmias' point can be illustrated as follows: I duck in a reflex because I see a ball coming towards me. Different regions of the brain are involved in the perception of the ball and my reflexive action. But it does not follow that the two are unconnected.

So it seems that Wegner's arguments for the illusory nature of conscious will are very weak indeed. So they hardly undermine free will, even if that requires the possibility of originating action by conscious decisions.

6. Concluding remarks

Clearly, what I've said is inconclusive. It is partly inconclusive because I have not discussed all relevant neuroscientific work (although I have discussed some of the most prominent work used in this debate). It is also inconclusive because I have not said very much about the requirements of free will itself. I have sympathy for the idea that it requires the ability to originate at least some of our actions consciously. But many philosophers disagree. Their conceptions of free will don't require this, and a more inclusive discussion would cover these as well.

Given my sympathy for the idea that free will requires the ability to originate actions consciously, I leave it open that neuroscience may one day provide good evidence against it. My point today is modest: the neuroscientific findings I've discussed do not provide strong evidence against free will. Given the

connection between free will and moral responsibility, they do not provide strong reason to rethink the foundations of our legal system.

References

Bayne, T. (2011), "*Libet and the Case for Free Will Scepticism*", in Swinburne, R. (ed.), *Free Will and Modern Science*, Oxford: Oxford University Press.

Bode, S., He, A., Soon, C., Trampel, R., Turner, R. & Haynes, J. (2011), "*Tracking the Unconscious Generation of Free Decisions Using Ultra-High Field fMRI*", PLoS ONE, Vol. 6: 1-13.

Dennett, D. & Marcel, K. (1992), "*Time and the Observer*", *Behavioral and Brain Sciences* 15(2), 183–247.

Haggard, P. & Eimer, M. (1999), "*On the Relation between Brain Potentials and the Awareness of Voluntary Movements*", *Experimental Brain Science*, Vol. 126: 128-133.

Libet, B., Gleason, C., Wright, E. & Pearl, D. (1983), "*Time of Unconscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential)*", *Brain*, Vol. 106: 623-642.

Libet, B. (1985), "*Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action*", *Behavioral and Brain Sciences*, Vol. 8: 529-566.

Mele, A. (2009), *Effective Intentions: the Power of Conscious Will*, Oxford: Oxford University Press.

Nahmias, E. (2002), "When Consciousness Matters: A Critical Review of Daniel Wegner's *The Illusion of Conscious Will*", *Philosophical Psychology*, Vol. 15: 527-541

Roediger, H., Goode M. & Zoromb F. (2008), "Free will and the control of action", in Baer, J., Kaufman, J. & Baumeister, R. *Are We Free? Psychology and Free Will*, New York: Oxford University Press.

Smith, K. (2011), "Neuroscience vs. Philosophy: Taking Aim at Free Will", *Nature news* website, <http://www.nature.com/news/2011/110831/full/477023a.html>.

Soon, C., Brass, M., Heinze, H. & Haynes, J. (2008), "Unconscious Determinants of Free Decisions in the Human Brain", *Nature Neuroscience*, Vol. 11: 543-545.

Wegner, D. (2002), *The Illusion of Conscious Will*, Cambridge: MA: MIT Press.