Do I Have Free Will?

Part II: The Challenge from Neuroscience

Jonathan Birch
j.birch2@lse.ac.uk

0 Recap of Part I

- To say that we have free will is to say that our actions are "up to us". This means that various alternative options are open to us, we can choose among them in accordance with our intentions and desires, and our actions originate in us and not in external causes.
- Free will is plausibly a necessary condition for moral responsibility. If an agent is not acting of their own free will, we don't hold them morally responsible for their actions.
- The historically most influential challenge to free will is the challenge from determinism, the idea that our actions are determined by facts beyond our control.
- One version of determinism is physical determinism, famously illustrated by Laplace's demon.
- Some interpretations of quantum physics reject physical determinism… but not in a way that makes any obvious room for free will.
- The challenge to free will from physical determinism is sharply formulated in the 'consequence argument'.

This time: Compatibilist escape routes… and the challenge from neuroscience

I Compatibilist escape routes

- A compatibilist responds to the challenge from determinism by asserting the compatibility of us having free will with our actions being determined by facts beyond our control.
- The consequence argument pushes the compatibilist into a tight corner.
- The compatibilist's best bet is to argue that we have been thinking about free will in the wrong way—we think it requires that alternative possibilities are open to us, but maybe it doesn't.

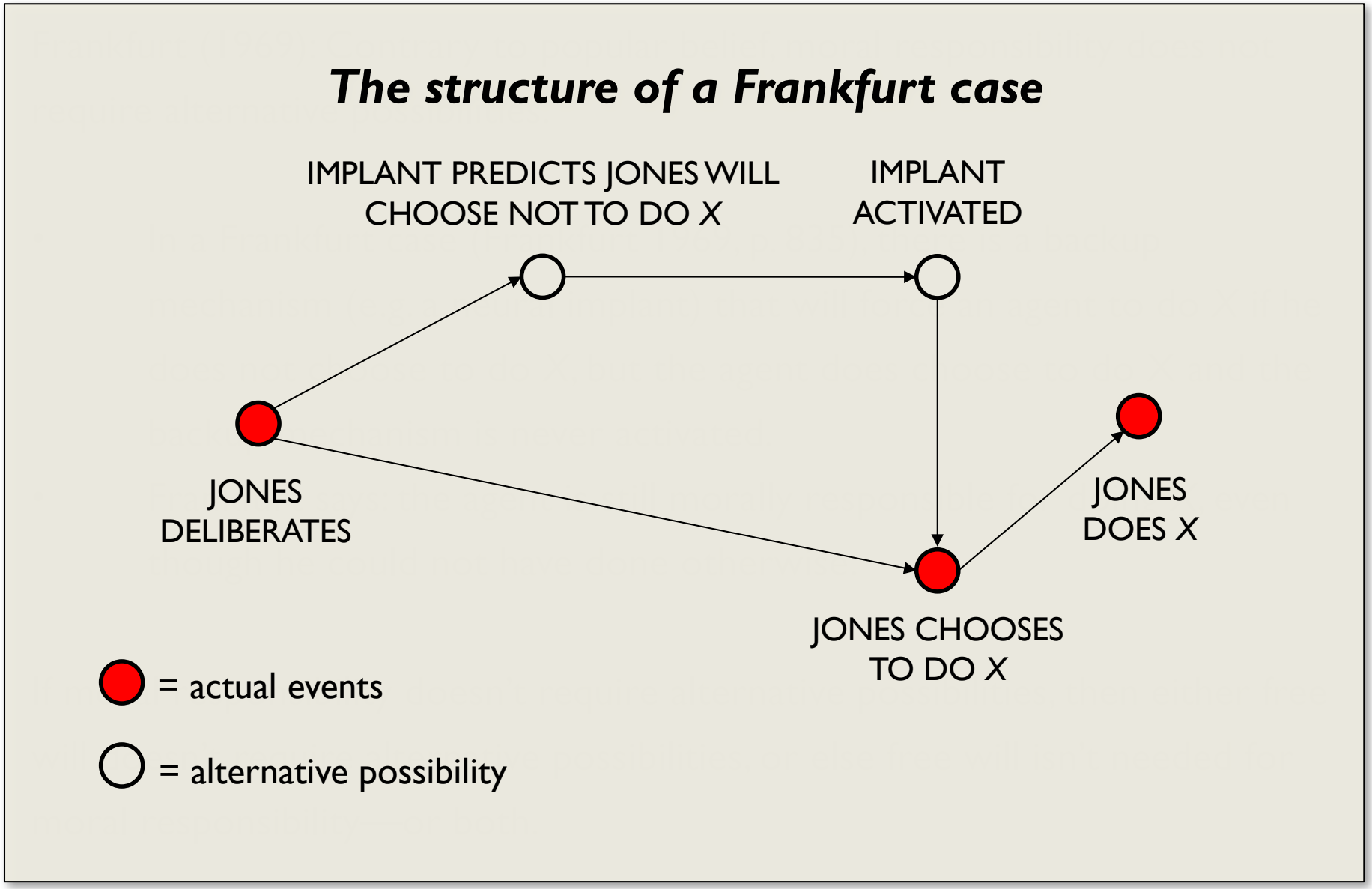
1 Compatibilist escape routes

Frankfurt (1969): Contrary to popular belief, moral responsibility does not require alternative possibilities.

- In a Frankfurt case (Frankfurt 1969, p. 835), there is a backup mechanism (e.g. a neural implant) that will force an agent to do *X* if it predicts he will not choose to do *X*, but the agent does choose to do *X* and the backup mechanism is never activated.
- Frankfurt says: the agent is still morally responsible for doing *X*, even though he could not have done otherwise.

If moral responsibility doesn't require alternative possibilities, then either free will doesn't require alternative possibilities, or else free will isn't needed for moral responsibility—or both.

Compatibilist escape routes



I Compatibilist escape routes

Frankfurt (1969): Contrary to popular belief, moral responsibility does not require alternative possibilities.

- In a Frankfurt case (Frankfurt 1969, p. 835), there is a backup mechanism (e.g. a neural implant) that will force an agent to do *X* if it predicts he will not choose to do *X*, but the agent does choose to do *X* and the backup mechanism is never activated.
- Frankfurt says: the agent is still morally responsible for doing *X*, even though he could not have done otherwise.

If moral responsibility doesn't require alternative possibilities, then either free will doesn't require alternative possibilities, or else free will isn't needed for moral responsibility—or both.

I Compatibilist escape routes

Frankfurt (1969): Contrary to popular belief, moral responsibility does not require alternative possibilities.

The "flicker of freedom" response

- In a Frankfurt case, the agent could have started forming an intention to not do *X*, even though this would have triggered the backup mechanism.
- This "flicker of freedom" at this initial stage explains why the agent is morally responsible for doing *X*.
- Determinism challenges even this flicker of freedom.

moral responsibility—or both.

1 Compatibilist escape routes

Frankfurt (1969): Contrary to popular belief, moral responsibility does not require alternative possibilities.

- In a Frankfurt case (Frankfurt 1969, p. 835), there is a backup mechanism (e.g. a neural implant) that will force an agent to do *X* if it predicts that he will not choose to do *X*, but the agent does choose to do *X* and the backup mechanism is never activated.
- Frankfurt says: the agent is still morally responsible for doing *X*, even though he could not have done otherwise.

If moral responsibility doesn't require alternative possibilities, then either free will doesn't require alternative possibilities, or else free will isn't needed for moral responsibility—or both.

2 Free will and second-order desires

Frankfurt's own view is that free will is the freedom to want what you want to want:

> "the statement that a person enjoys freedom of the will means . . . that he is free to want what he wants to want. More precisely, it means that he is free to will what he wants to will, or to have the will that he wants. Just as the question about the freedom of an agent's action has to do with whether it is the action he wants to perform, so the question about the freedom of the will has to do with whether it is the will that he wants to have." (Frankfurt 1971, p. 15)

2 Free will and second-order desires

- First-order desires = desires for ordinary objects.
- Second-order desires = desires you have about which first-order desires you want to motivate you.

Free will for Frankfurt is a special type of alignment between your first-order desires and your second-order desires.

Advantages:

- Explains the difference between willing and unwilling/wanton addicts.
- Identifies an important difference between human agency and agency in (most?) non-human animals.
- Makes free will "worth wanting"

… but does it really escape the challenge from determinism?

A problem for Frankfurt

Is it true that an agent with free will could have willed otherwise?

If yes, how is this compatible with determinism?

If no, in what sense is the agent "free to will what he wants to will"?

Frankfurt wants it both ways…

> Whatever his will, then, the will of the person whose will is free could have been otherwise; he could have done otherwise than to constitute his will as he did.

p. 19

> My conception of the freedom of the will appears to be neutral with regard to the problem of determinism.

p. 20

3 The challenge from neuroscience

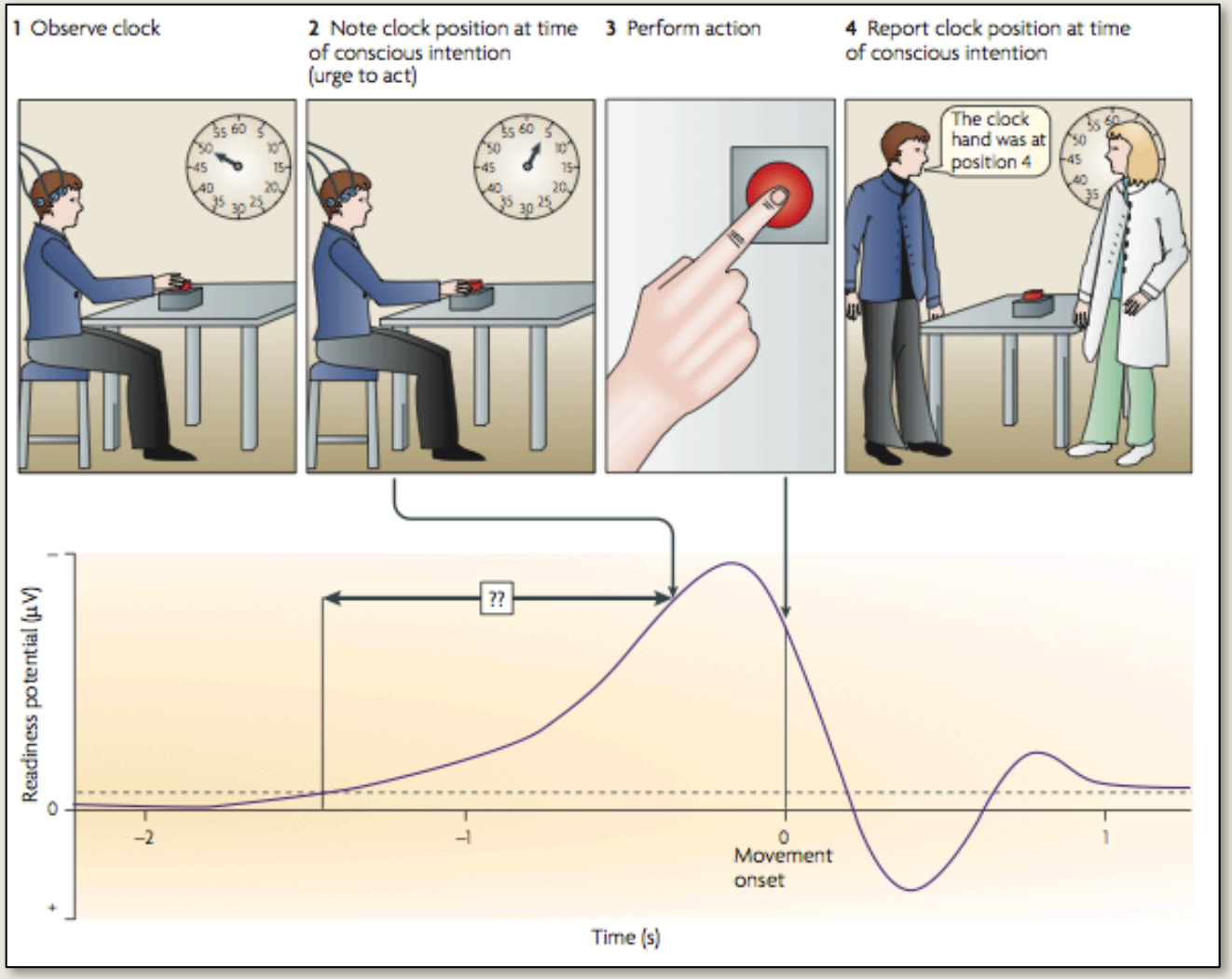
Compatibilist theories of free will often throw out the alternative possibilities and control components of free will as traditionally understood, but they still hold on to a version of the origin component:

> An action is freely willed only if it is initiated by the agent's conscious choices, desires and intentions.

Does this have to go too? Recent work in the neuroscience of agency suggests it might.

The 'Libet experiments'

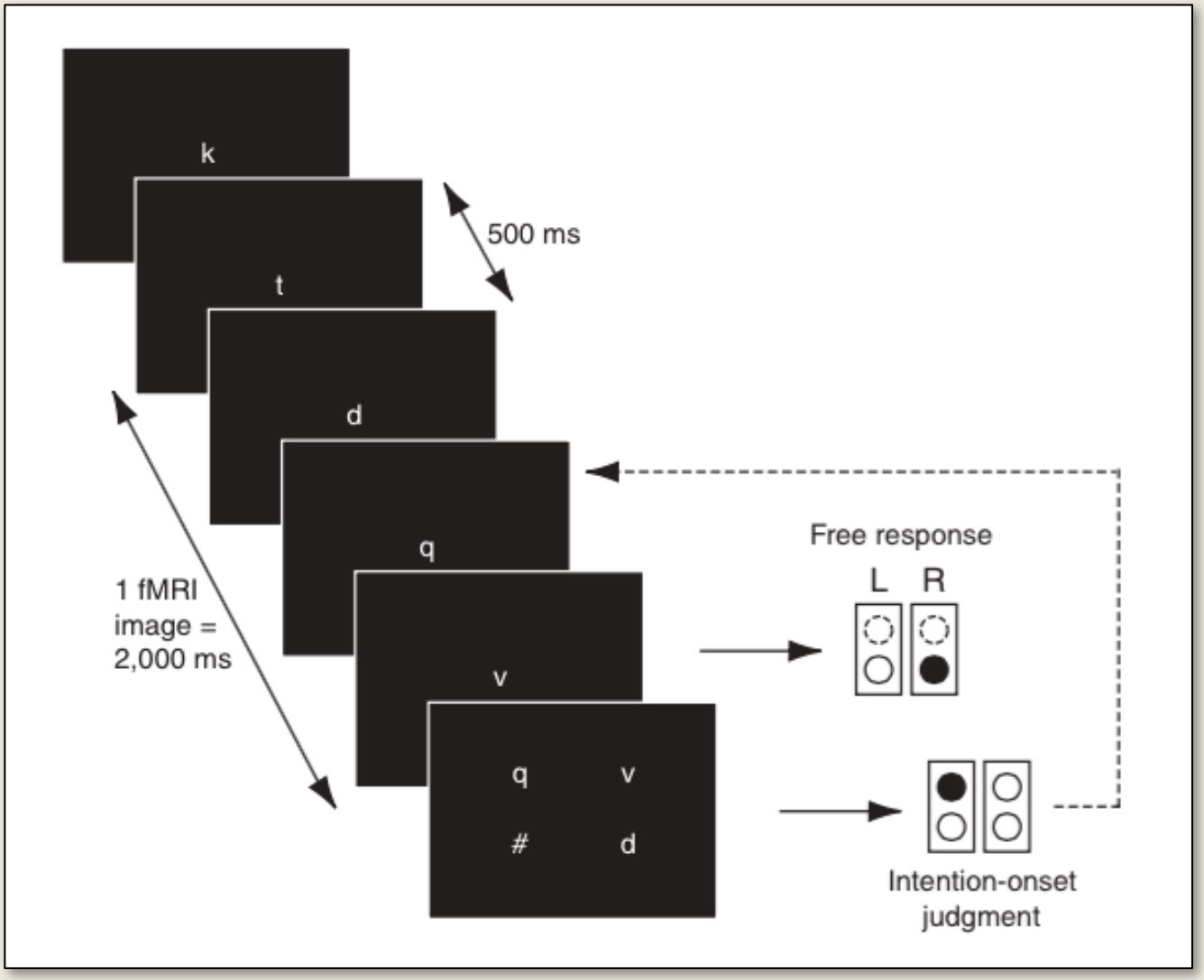
from Haggard 2008



3 The challenge from neuroscience

Soon and colleagues' update

from Soon et al. 2008



3 The challenge from neuroscience

What is the nature of the threat posed by the Libet experiments?

Weak version:

Conscious choices, desires and intentions do not initiate action. They occur after an action has been unconsciously initiated.

Strong version:

Conscious choices, desires and intentions are not even causally relevant to action. They are wholly epiphenomenal (see Week 4).

Even the weak version of the challenge creates a serious problem for free will.

4 Free will saved?

Libet:

The results can be reconciled with free will, because the subject can consciously choose to 'veto' the action after the readiness potential has started.

> Potentially available to the conscious function is the possibility of stopping or vetoing the final progress of the volitional process, so that no actual muscle action ensues. *Conscious-will could thus affect the outcome* of the volitional process even though the latter was initiated by unconscious cerebral processes. Conscious-will might block or veto the process, so that no act occurs.

> All of us, not just experimental subjects, have experienced our vetoing a spontaneous urge to perform some act. This often occurs when the urge to act involves some socially unacceptable consequence, like an urge to shout some obscenity at the professor.

4 Free will saved?

Libet:

The results can be reconciled with free will, because the subject can consciously choose to 'veto' the action after the readiness potential has started.

> Potentially available to the conscious function is the possibility of stopping or vetoing the final progress of the volitional process, so that no actual muscle action ensues. *Conscious-will could thus affect the outcome* of the volitional process even though the latter was initiated by unconscious cerebral processes. Conscious-will might block or veto the process, so that no act occurs.

However…

- This seems like 'free will' only in a very attenuated sense.
- Libet's results don't rule out that these 'conscious vetoes' are also initiated by unconscious neural activity.

4 Free will saved?

Libet:

The results can be reconciled with free will, because the subject can consciously choose to 'veto' the action after the readiness potential has started.

> One should, at this point, consider the possibility that the conscious veto itself may have its origin in preceding unconscious processes, just as is the case for the development and appearance of the conscious will. If the veto itself were to be initiated and developed unconsciously, the choice to veto would then become an unconscious choice of which we *become* conscious, rather than a consciously causal event. Our

However…

- This seems like 'free will' only in a very attenuated sense.
- Libet's results don't rule out that these 'conscious vetoes' are also initiated by unconscious neural activity.

4 Free will saved?

Libet:

The results can be reconciled with free will, because the subject can consciously choose to 'veto' the action after the readiness potential has started.

Mele:

Libet asked subjects to press the button when they 'felt the urge' to do so. One interpretation of the readiness potential is that it reflects this 'urge' arising. But since an urge is not the same thing as an intention to act on the urge, the results do not present a serious challenge to free will.

4 Free will saved?

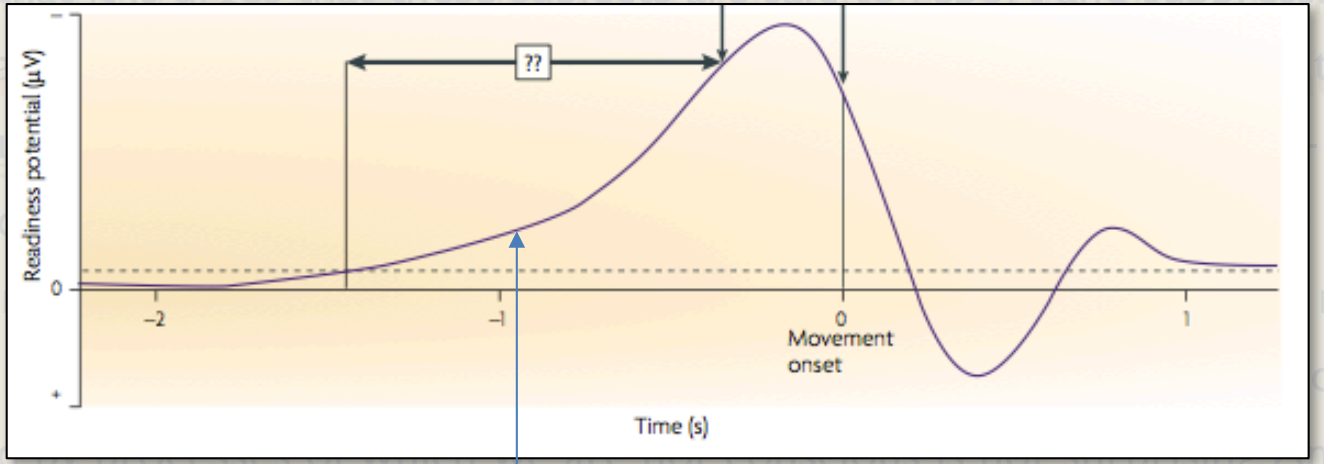
Mele's critique of Libet

"Notice that it is urges that these subjects are said to report and suppress. Might it be that [activity in the first 300ms] is a potential cause of conscious urges to flex in Libet's subjects and some subjects make no decision about when to flex—unconsciously or otherwise—until after the conscious urge emerges? And might it be that prior to the emergence of the conscious urge, these subjects have no proximal intention to flex—not even an unconscious one? That our urges often are generated by processes of which we are not conscious is not surprising. And if we sometimes make effective decisions about whether or not to act on a conscious urge, so much the better for free will."

Mele (2011), p. 503.

4 Free will saved?

Mele's critique of Libet



Is this just an urge arising?

4 Free will saved?

Mele's critique of Libet

"Notice that it is urges that these subjects are said to report and suppress. Might it be that [activity in the first 300ms] is a potential cause of conscious urges to flex in Libet's subjects and some subjects make no decision about when to flex—unconsciously or otherwise—until after the conscious urge emerges? And might it be that prior to the emergence of the conscious urge, these subjects have no proximal intention to flex—not even an unconscious one? That our urges often are generated by processes of which we are not conscious is not surprising. And if we sometimes make effective decisions about whether or not to act on a conscious urge, so much the better for free will."

Mele (2011), p. 503.

5 Summary

- To escape the challenge from determinism, compatibilists tend to argue that we have been thinking about free will in the wrong way—we think it requires that alternative possibilities are open to us, but it doesn't.
- Frankfurt cases, in which a backup mechanism ensures that an agent cannot do anything other than what they actually do, are used to motivate this idea.
- An example of an (allegedly) compatibilist theory of free will is Frankfurt's theory, which says that free will is the freedom to want what you want to want. Whether this theory really escapes the challenge from determinism is debatable.
- Even compatibilist theories of free will usually assume that an action is freely willed only if initiated by the agent's conscious choices, desires and intentions.
- However, the Libet experiments (and recent updates) suggest that actions are initiated by unconscious brain processes that begin 0.4-10s before a conscious intention is formed.
- Libet himself thought this left room for a 'conscious veto' or 'free won't'.
- Mele argues that Libet conflates urges with intentions, casting doubt on the significance of his work for the free will debate.

For more on this topic: Take PH221, Problems of Analytic Philosophy.