

Frankfurt
examples

the
Libet
experiments

challenging
Libet

free

will

vs.

neuroscience

In ancient times, philosophers worried about the challenge to free will from fate. After the birth of modern physics, many worried about the challenge to free will from determinism. Our topic today is a much more recent challenge: the challenge to free will posed by contemporary [neuroscience](#).

Before we get into that topic, though, I want to discuss an important challenge to some of the arguments we have discussed so far.

The consequence argument for the incompatibility of free will and determinism, the fatalist argument, and the foreknowledge argument all had the following premise in common:

The no choice principle

If no one has any choice about whether P, and no one has any choice about whether, if P, then Q, then no one has any choice about whether Q

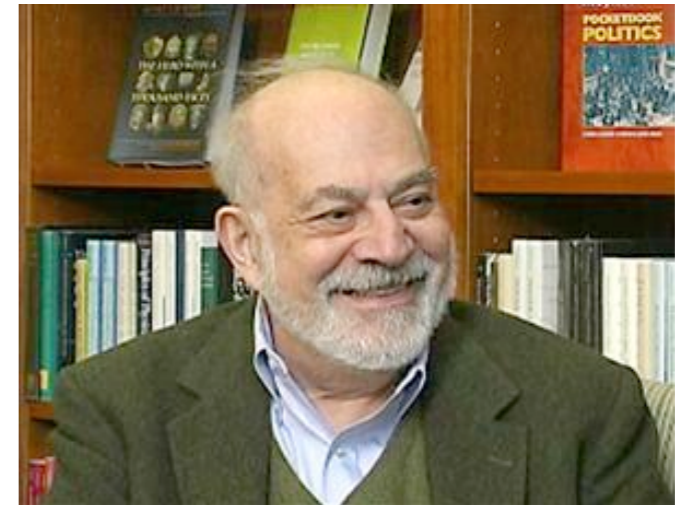
This assumption seems, for reasons we have discussed, very plausible. But it can be challenged.

The no choice principle

If no one has any choice about whether P, and no one has any choice about whether, if P, then Q, then no one has any choice about whether Q

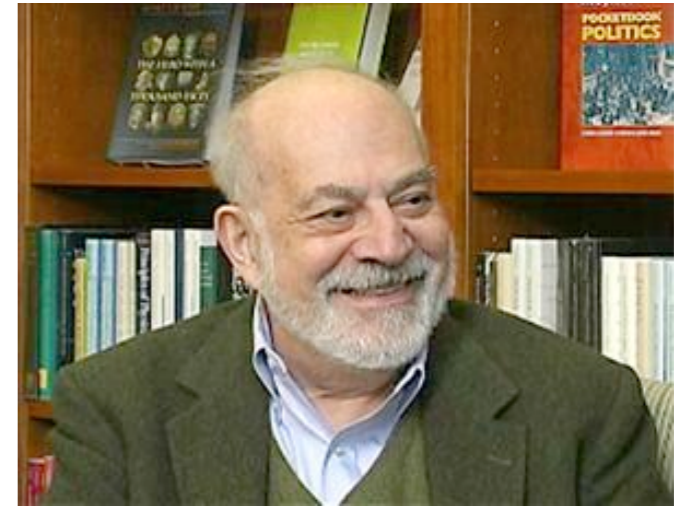
This assumption seems, for reasons we have discussed, very plausible. But it can be challenged.

These are due to the contemporary philosopher Harry Frankfurt.



Here's a more recent version (due to John Martin Fischer) of the kind of example Frankfurt used to challenge the no choice principle.

Here's a more recent version (due to John Martin Fischer) of the kind of example Frankfurt used to challenge the no choice principle.



“Imagine, if you will, that Black is a quite nifty (and even generally nice) neurosurgeon. But in performing an operation on Jones to remove a brain tumor, Black inserts a mechanism into Jones’s brain which enables Black to monitor and control Jones’s activities. Jones, meanwhile, knows nothing of this. Black exercises this control through a sophisticated computer which he has programmed so that, among other things, it monitors Jones’s voting behavior. If Jones were to show any inclination to vote for Bush, then the computer, through the mechanism in Jones’s brain, intervenes to ensure that he actually decides to vote for Clinton and does so vote. But if Jones decides on his own to vote for Clinton, the computer does nothing but continue to monitor - without affecting the goings-on - in Jones’s head.”

“Imagine, if you will, that Black is a quite nifty (and even generally nice) neurosurgeon. But in performing an operation on Jones to remove a brain tumor, Black inserts a mechanism into Jones’s brain which enables Black to monitor and control Jones’s activities. Jones, meanwhile, knows nothing of this. Black exercises this control through a sophisticated computer which he has programmed so that, among other things, it monitors Jones’s voting behavior. If Jones were to show any inclination to vote for Bush, then the computer, through the mechanism in Jones’s brain, intervenes to ensure that he actually decides to vote for Clinton and does so vote. But if Jones decides on his own to vote for Clinton, the computer does nothing but continue to monitor - without affecting the goings-on - in Jones’s head.”

Suppose now that Jones decides “on his own” to vote for Clinton. In this case the mechanism which Black has inserted in Jones’ brain never kicks on and forces Black to do anything.

In this situation, is Black’s act of voting for Clinton free?

You might think: yes. After all, Black decided to do this, and nothing forced him to do so.

Suppose now that Jones decides “on his own” to vote for Clinton. In this case the mechanism which Black has inserted in Jones’ brain never kicks on and forces Black to do anything.

In this situation, is Black’s act of voting for Clinton free?

You might think: yes. After all, Black decided to do this, and nothing forced him to do so.

But now consider the following points:

Jones had no choice about the fact that Black inserted the mechanism into his brain.

Jones had no choice about the fact that if Black inserted the mechanism in his brain, he would vote for Clinton.

Despite these facts, it seems that Jones had a choice about whether to vote for Clinton.

Jones had no choice about the fact that Black inserted the mechanism into his brain.

Jones had no choice about the fact that if Black inserted the mechanism in his brain, he would vote for Clinton.

Despite these facts, it seems that Jones had a choice about whether to vote for Clinton.

If you accept all of these claims, this looks like a counterexample to the no choice principle.

More generally, it looks like an example of an action in which the agent was determined to do something by past facts, and was nonetheless free. So it looks like an argument that free will is compatible with determinism.

It also gives us the resources to reject one of the premises of the fatalist and foreknowledge arguments, and thereby show that free will is compatible both with fate and with divine foreknowledge of our actions.

More generally, it looks like an example of an action in which the agent was determined to do something by past facts, and was nonetheless free. So it looks like an argument that free will is compatible with determinism.

It also gives us the resources to reject one of the premises of the fatalist and foreknowledge arguments, and thereby show that free will is compatible both with fate and with divine foreknowledge of our actions.

This all looks like good news for free will. But there are two things worth keeping in mind.

First, recall that a compatibilist view of free will seems to undercut the free will defense. For if free will were compatible with determinism, then (as Mackie says) God could have given us free will while also determining that we would never bring about any evil. So if you think that the free will defense explains the existence of at least some evil, you can't accept the Black/Jones example at face value.

First, recall that a compatibilist view of free will seems to undercut the free will defense. For if free will were compatible with determinism, then (as Mackie says) God could have given us free will while also determining that we would never bring about any evil. So if you think that the free will defense explains the existence of at least some evil, you can't accept the Black/Jones example at face value.

Second, it is not completely obvious that this example does show that free will is compatible with determinism.

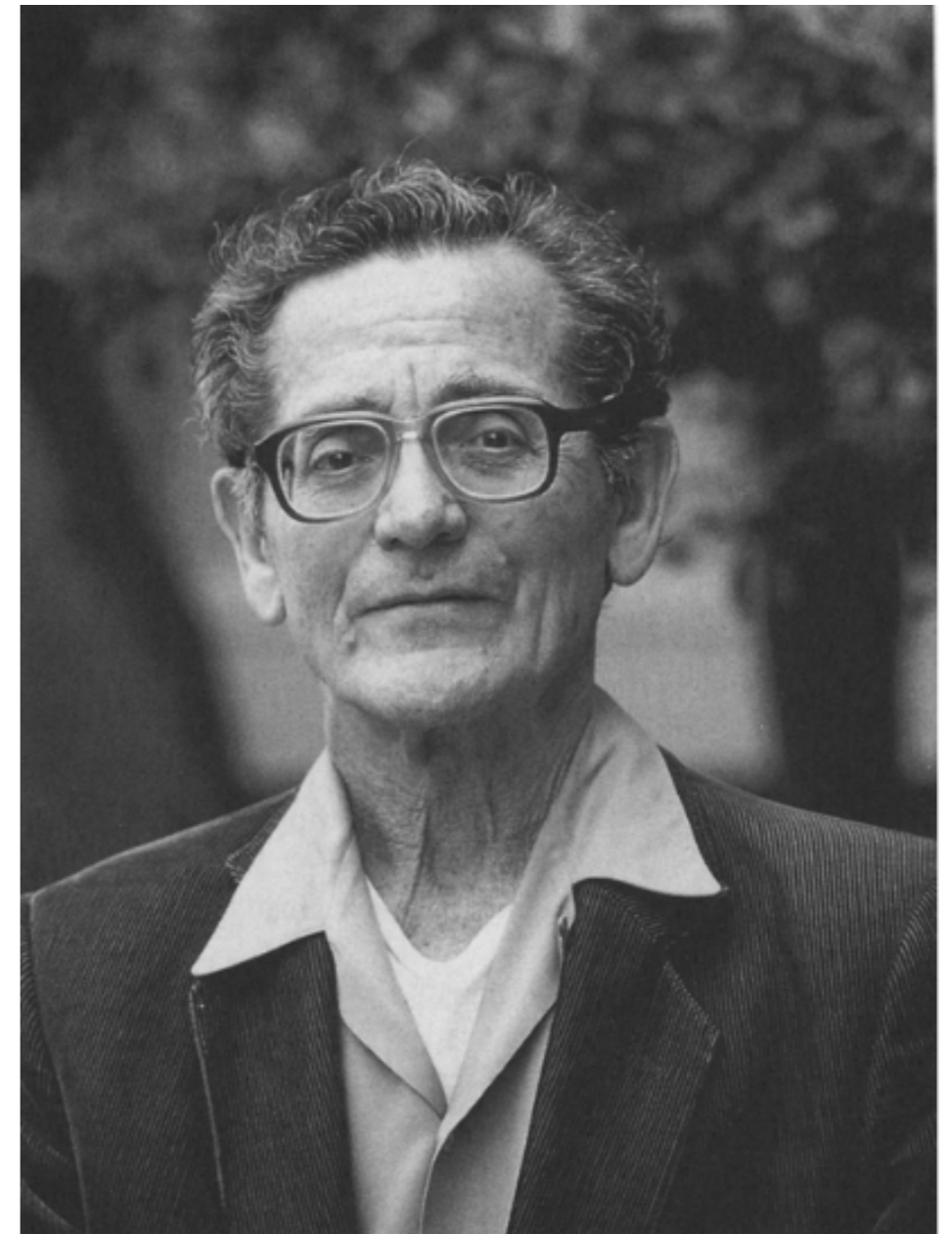
Let's distinguish between the act of Jones deciding at a certain moment to vote for Clinton, and the act of Jones actually voting for Clinton. There is at least some tendency to say that the act of deciding was free, but that the act of voting was not (since the latter was determined to happen by the mechanism inserted in Jones' brain).

If you think this, then the case is not a counterexample to the no choice principle (since the mechanism did not determine that Jones would make the choice to decide to vote for Clinton at that time).

Let's turn now to the challenge to free will posed by contemporary neuroscience.

In particular, our focus will be on some groundbreaking experimental results obtained by the late American neuroscientist Benjamin Libet.

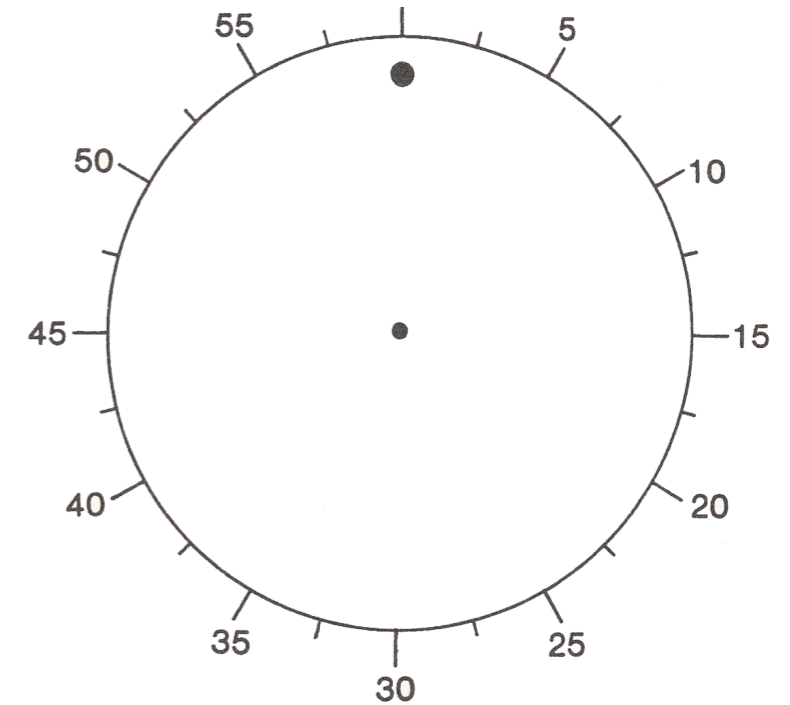
Libet's work was on the neuroscience of consciousness. Since Libet thought, not unreasonably, that free choices had to be conscious, he thought that we could try to design experiments which would show whether or not people had free will.



In the central experiment described in the reading for today, subjects were told to look at a clock with a dot which moved rapidly in circles around the clock.

In the central experiment described in the reading for today, subjects were told to look at a clock with a dot which moved rapidly in circles around the clock.

Here is how Libet describes the instructions given to these subjects:



The subject was asked to wait for one complete revolution of the CRO spot and then, at any time thereafter when he felt like doing so, to perform the quick, abrupt flexion of the fingers and/or the wrist of his right hand (*see Libet et al., 1982*). An additional instruction to encourage 'spontaneity' of the act was given routinely to subjects in Group 2 and only in the latter half to two-thirds of sessions with Group 1. For this, the subject was instructed 'to let the urge to act appear on its own at any time without any preplanning or concentration on when to act', that is, to try to be 'spontaneous' in deciding when to perform each act; this instruction was designed to elicit voluntary acts that were freely capricious in origin.

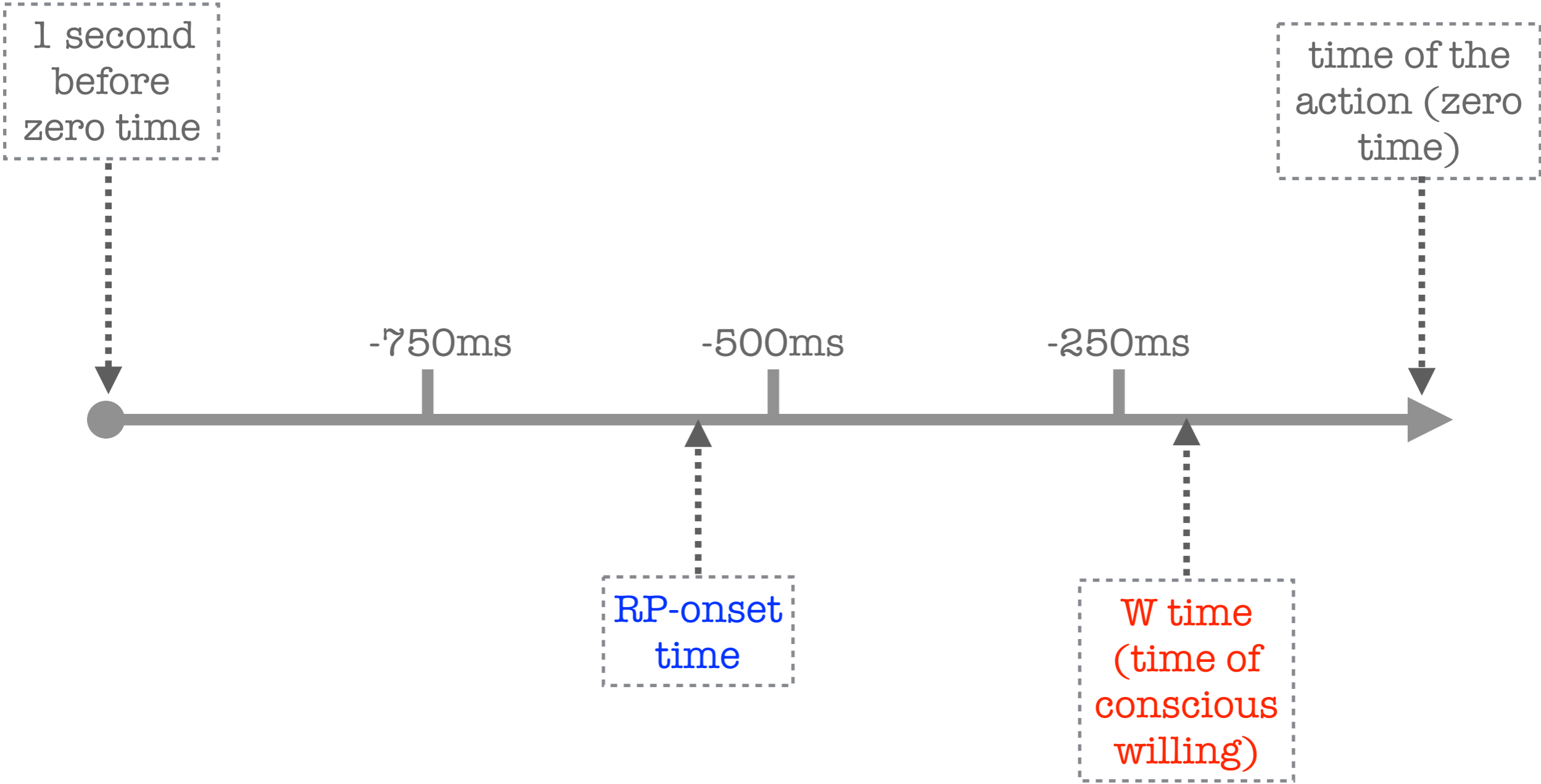
Subjects were then asked to note where the spot on the clock was when they had the urge, or desire, to flex. This was used to record the time of, as Libet thought of it, the subject's **conscious willing** to flex his or her hand. Libet called this the "W time."

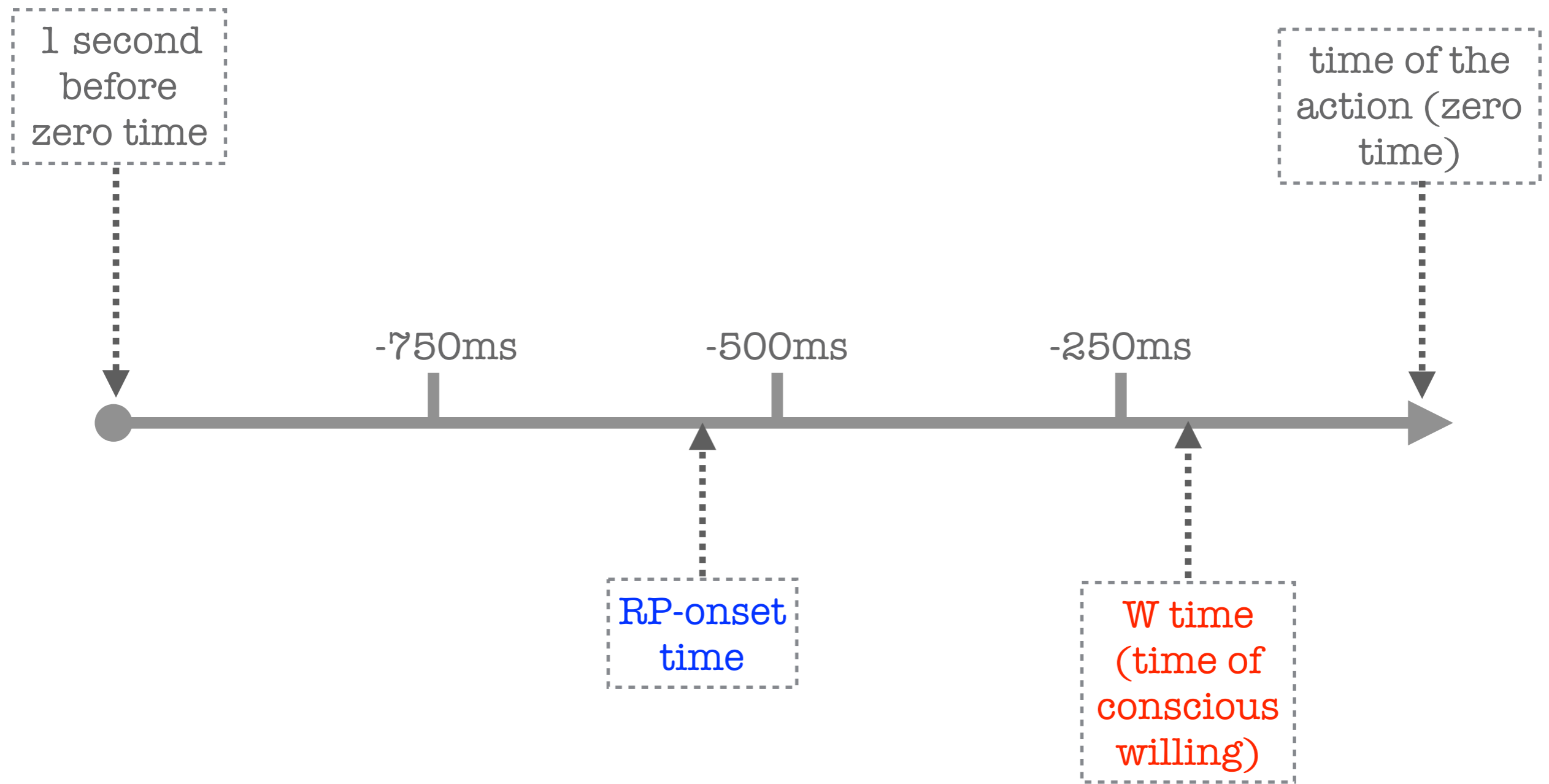
Subjects were then asked to note where the spot on the clock was when they had the urge, or desire, to flex. This was used to record the time of, as Libet thought of it, the subject's **conscious willing** to flex his or her hand. Libet called this the "**W time**."

This was then compared with the time at which certain brain events, measured by EEG, occurred in the subject. These were brain events which other experiments had shown to precede certain intentional actions. The increased brain activity which occurs prior to a certain sort of intentional action is called that action's **readiness potential**. Libet called the times at which subjects showed a readiness potential for flexing their hands "**RP-onset times**."

The W time and the RP-onset time were then compared with the "zero time" — the time at which the subject's hand actually flexed.

When Libet compared these times, he found something remarkable.





It appears that the subject's brain is ready to flex the hand about 350ms **before** the subject's experience of consciously deciding to flex his or her hand.

This makes it seem as though the conscious "decision" to flex one's hand is not really a decision at all — that decision has already been made, unconsciously, by the brain.

It appears that the subject's brain is ready to flex the hand about 350ms **before** the subject's experience of consciously deciding to flex his or her hand.

This makes it seem as though the conscious "decision" to flex one's hand is not really a decision at all — that decision has already been made, unconsciously, by the brain.

And, Libet thought, an unconscious decision made in the brain, prior to any conscious act of deciding, cannot be free; free decisions must be consciously made.

Does this show that there is **no** space for conscious free will? Libet thought not.

Does this show that there is **no** space for conscious free will? Libet thought not.

There could be a conscious ‘veto’ that aborts the performance even of the type of ‘spontaneous’ self-initiated act under study here. This remains possible because reportable conscious intention, even though it appeared distinctly later than onset of RP, did appear a substantial time (about 150 to 200 ms) before the beginning of the movement as signalled by the EMG.

Libet thought that, in the time between W time and the time of the action, the subject may be able to block the execution of the action which had already been decided on, unconsciously, by the brain.

And in fact Libet carried out further experiments which he took to show that this is indeed possible.

And in fact Libet carried out further experiments which he took to show that this is indeed possible.

In these experiments, subjects were instructed to do two things. (1) Prepare to flex at a specific **target time** — say, when the dot is at “30.” (2) Do not flex at that time.

In these experiments, Libet observed higher EEG readings — and thus readiness potential to flex — about a second before the target time. These EEG readings were remarkably similar to those at about -500ms in the original study.

However, the EEG readings decreased around 200ms before the target time — not far off of the W time from the previous experiment.

Libet took this to mean that the subjects in the “veto” experiment decided to flex at the target time, but were able to exercise conscious free will to veto this decision about 200ms before the action.

Libet took this to mean that the subjects in the “veto” experiment decided to flex at the target time, but were able to exercise conscious free will to veto this decision about 200ms before the action.

Is this good news for free will? Yes and No. Yes, because it appears to make room for conscious free will. No, because it gives conscious free will a disappointingly limited role to play.

In another paper, Libet described the situation as follows:

Assuming that one can extrapolate these results to volitional acts generally, they do not exclude a possible role for free will, even though the volitional process starts with unconscious cerebral activity. However, the potential role of free will would be constrained; it would be changed from being an initiator of the voluntary act to one only of controlling the outcome of the volitional process, after the individual becomes aware of an intention or wish to act now. In a general sense, free will could only select from among the brain activities that are a part of a given individual’s makeup.

From Libet et. al., “The Neural Time-Factor in Perception, Volition, and Free Will”

The Libet experiments are a nice example of the interconnectedness of science and philosophy. Often in the history of philosophy, philosophers have formulated a deep and interesting question, which then inspired scientists (who, in many cases, were themselves philosophers) to formulate experiments which promised to answer the question.

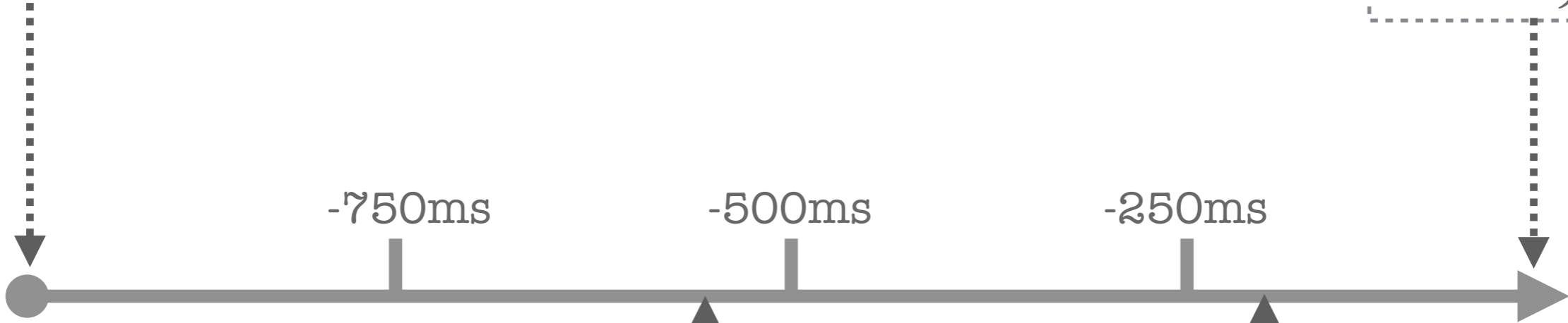
Our question is: do Libet's experiments show that free will is limited in the way that he suggests?



I want to discuss two ways of challenging Libet's conclusion.

1 second before zero time

time of the action (zero time)



RP-onset time

W time (time of conscious willing)

In his experiments, Libet took the EEG readings at RP-onset time to indicate an unconscious decision to act. This leads to two questions.

Is RP-onset an unconscious decision?

If RP-onset is a decision, does that limit our freedom?

Frankfurt examples

the Libet experiments

challenging Libet

Is RP-onset an
unconscious
decision?

There is wide experimental confirmation of the fact that the sort of increased brain activity which occurs at RP-onset is correlated with actions. But that does not mean that it is an unconscious decision. Perhaps, for instance, RP-onset is a process which sometimes leads to a decision, rather than the decision itself. Maybe it just shows that the action is being considered, or imagined.

Some aspects of Libet's experiments, in fact, suggest that RP-onset is not a decision.

Is RP-onset an
unconscious
decision?

Recall the “veto” experiment, in which subjects were asked to prepare to flex their hands at a certain time, but then **not** flex them at that time.

In that case, the electrical activity in the brain was extremely similar to that observed at RP-onset in the original experiment.

But did subjects in the veto experiment ever decide to flex their hands?

Suppose that I asked you to prepare to sing the Fight Song in 2 minutes, but not do it. Would you have decided to sing the Fight Song?

Indeed, it seems impossible to decide to do something that you have also decided not to do. If I offered you a large reward to for deciding, at will, to sing the Fight Song and then not do it, you would not be able to claim the reward.

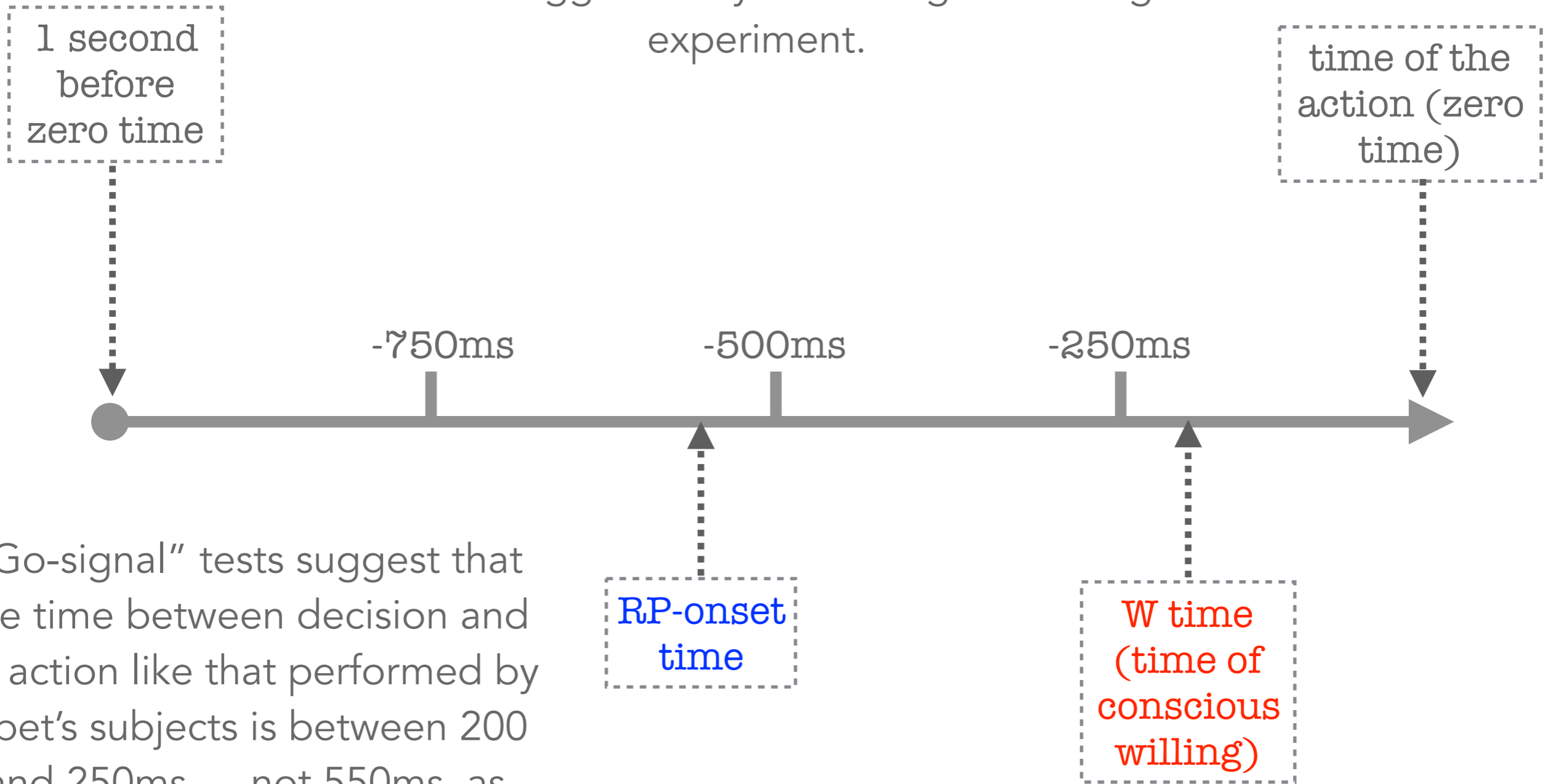
Is RP-onset an
unconscious
decision?

But did subjects in the veto experiment ever decide to flex their hands?

If subjects in the “veto” experiment never decided to flex their hands, and their brain activity was very similar to that observed at RP-onset, that strongly suggests that RP-onset is not a conscious decision.

This is also suggested by the timing of the original experiment.

This is also suggested by the timing of the original experiment.



“Go-signal” tests suggest that the time between decision and an action like that performed by Libet’s subjects is between 200 and 250ms — not 550ms, as would be the case if RP-onset were a decision.

This also fits nicely with the alternative hypothesis that the decision does not take place at RP-onset time, but at W time.

If RP-onset is a decision, does that limit our freedom?

These are serious worries about Libet's argument. But let's set them aside and ask: if RP-onset is an unconscious decision, how damaging is that to our belief in freedom of the will?

One might think that it is not very damaging, on the grounds that the actions subjects perform in the Libet experiments are in some ways different than paradigm examples of free action.

Libet seems to disagree with this:

The present evidence for the unconscious initiation of a voluntary act of course applies to one very limited form of such acts. However, the simple voluntary motor act studied here has in fact often been regarded as an incontrovertible and ideal example of a fully endogenous and 'freely voluntary' act. The absence of any larger meaning in the simple quick flexion of hand or fingers, and the possibility of performing it with capriciously whimsical timings, appear to exclude external psychological or other factors as controlling agents. It thus invites the extrapolation that other relatively 'spontaneous' voluntary acts, performed without conscious deliberation or planning, may also be initiated by cerebral activities proceeding unconsciously.

If RP-onset is a decision, does that limit our freedom?

This is not unreasonable — we do often use simple acts, like deciding to scratch one's nose, as an example of a free action.

But Libet's subjects are in one central respect different than subjects of ordinary free actions: they are asked to be as spontaneous as possible, and avoid planning when they will flex their fingers.

That is not how most free actions work; in the case of most free actions, we consciously consider pros and cons of the action, and plan when to carry the action out.

This fact leads to a possibility that is worth considering.

This fact leads to a possibility that is worth considering.

Consider your decision to come to Notre Dame. Here is one way things could have gone:

The decision

You consciously thought about it for a long time. You weighed the pros (great academic reputation, football, wonderful philosophy professors) and the cons (weather, the university theology requirement). Finally, after months of stewing, you decided to come to Notre Dame. That decision involved a brain event and a conscious awareness of the decision. The brain event (RP-onset) occurred about 300ms before the conscious awareness.

Would the fact that that the brain event occurred 300ms before the conscious awareness make your decision unfree? If not, then why should we think that Libet's data — even if RP-onset is an unconscious decision — tells us much about free will?

Is RP-onset an
unconscious
decision?

If RP-onset is a
decision, does
that limit our
freedom?

Libet's experiments are fascinating. But one can challenge both his views about when unconscious decisions occur, and his views about the significance of the timing of unconscious decisions.

There is much ongoing work in neuroscience and social science about the will and freedom of the will. If you'd like to know more, a good overview is the philosopher Alfred Mele's book *Free*. Many of the critical points made above are due to his work.