

# What does it mean to inhibit an Action?

## A Critical Discussion of Benjamin Libet's Veto in a Recent Study

Robert Reimer<sup>[0000-0003-0947-8249]</sup>

Universität Leipzig, Fakultät für Philosophie,  
Beethovenstraße 15, 04107 Leipzig, Germany  
rreimer@posteo.de

**Abstract:** In the 1980s, physiologist Benjamin Libet conducted a series of experiments to test whether the will is free. Whilst he originally assumed that the will functions like an immaterial initiator of cerebral processes culminating in actions, he later began to think that it rather works like an immaterial veto inhibiting unwanted actions by preventing unconsciously initiated cerebral processes from unfolding. Libet's veto was widely criticized for its Cartesian dualist and interactionist implications. However, in 2016, Schultze-Kraft et al. adopted Libet's idea of an action-inhibiting veto and conducted a new experiment. Its goal was to test until which moment agents can inhibit an action that they already intended to do. Despite insisting on the material nature of the veto, the researchers also described the function of the veto in interactionist terms, namely as an act of the agent performed against her own cerebral processes. The purpose of this paper is to explain in which sense the veto in Libet's and in Schultze-Kraft's study is interactionist, and to provide a non-interactionist reinterpretation of the test subjects' action inhibition in Schultze-Kraft's experiment.

**Keywords:** B. Libet, R. Descartes, Veto, Action Inhibition, Interactionism

### 1 Introduction

In the 1980s, physiologist Benjamin Libet conducted a series of experiments to prove that the will is free. He asked his test subjects to flex their wrists at a moment of their own choice. Before acting, the test subjects should determine the point in time when they felt the urge to perform this movement. After a series of trials, Libet found out that the awareness of the urge was preceded by an unconscious, slow electrical change recordable on the scalp at the vertex by 300 ms or more (Libet 1999, 49). This 'readiness potential' (RP) was already discovered by Kornhuber & Deecke (1965) and was associated with the performed hand movement. Libet concluded that the action of the test subjects did not originate in the test subjects' urge but rather in the RP.

Disappointed by these results, Libet considered another option of how the freedom of the will could be saved. In some of his early experiments, test subjects reported that they were able to suppress the urge to act before any actual movement occurred (Libet

1983, 641). Based on this observation, Libet argued that the will rather functions like a veto. Instead of being the initiator of the action, it is able to interfere with or control the unconscious processes triggered by the original RP by aborting or selecting some of them consciously before they culminate in bodily movement. This assumption was supported by the results of one of his later experiments, in which the test subjects were explicitly instructed to veto the development of their action (idem. 1985, 538). Assuming that these mental acts of veto are not themselves caused by any unconscious cerebral activities, Libet concluded that the will can indeed be free. He noted:

“Although the volitional process may be initiated by unconscious cerebral activities, conscious control of the actual motor performance of voluntary acts definitely remains possible. The findings should therefore be taken not as being antagonistic to free will but rather as affecting the view of how free will might operate. Processes associated with individual responsibility and free will would ‘operate’ not to initiate a voluntary act but to select and control volitional outcomes.” (ibid.)

Many criticized Libet’s notion of an action inhibiting veto because it is based on an implausible *dualist interactionist paradigm*. Dualist interactionism is a theory of the mind according to which mind and brain are two distinct substances whose internal processes, in general, unfold independently from one another. However, the mind can still influence the processes of the brain by causing an effect in one of its parts.

Despite this criticism, Libet’s idea of the veto was adapted and tested in one way or another by scientists in some recent studies. Schultze-Kraft et al., for instance, designed an experimental setup in which test subjects should perform a simple foot movement while facing a green light. If the light turned red, however, they were supposed to inhibit their action by ‘exerting a veto’ against the already prepared and upcoming movement (Schultze-Kraft et al. 2016, 1080). With their study, Schultze-Kraft et al. tried to determine the point in time after which people can no longer inhibit their action.

The purpose of this paper is to show that, despite explicitly rejecting Libet’s dualist interpretation of the veto, Schultze-Kraft’s study still falls prey to the interactionism that is underlying Libet’s interpretation. In section two, I will specify the general form of Libet’s veto and discuss in which sense it is an expression of a dualist interactionist view. I will also say a few words on the historical background of Libet’s idea and show that it can already be found in René Descartes’ account on the body-soul interaction. In section three, I will pin down the interactionist character of the veto in Schultze-Kraft’s interpretation and criticize it. In section four, I will then suggest a more plausible and non-interactionist interpretation of what actually happened when the test subjects inhibited their action in Schultze-Kraft’s study.

## 2 Descartes, Libet, and the Veto

Any kind of interactionism presupposes the distinction between two distinct and mutually independent systems *in* a person. These are commonly mind and brain, or soul and body, but, as we will see later, they can also be the person herself and her brain. In each

of these systems, many processes take place independently from the processes of the other system. However, one system can interfere with the processes of the other.<sup>1</sup>

Philosophers often identify Descartes as the ‘founding father’ of interactionism. In ‘The Passions of the Soul’, Descartes develops an account on action and perception that is consistent both with his own biological research and his religious faith. Against this background, Descartes posits the existence of two different substances constituting a person (an agent) – soul and body. The soul (the first system) is an immaterial and immortal substance that thinks and calculates. It is connected with the body through the pineal gland in the brain (Descartes 2014 art. 32). The body (the second system) is a material and mortal substance, in turn. Various material processes such as digestion and limb movement take place in the body. But whilst the soul is absolutely free from the influences of the body (ibid. art. 41), agents can feel and move their bodies *through* their souls (ibid. 17) by performing an act of will. The act of will (mediated by the soul and operating through the pineal gland) causes the nerves in the brain to ‘vibrate’ and then the muscles in the rest of the body to contract (ibid., art. 34). This account on action is *interactionist* because it suggests that the way how the agent acts is by interfering with the processes of her body (system 1) mediated through her soul (system 2).

The standard function of the will, according to Descartes, is the initiation of cerebral processes. However, the agent can also abort an unfolding cerebral process by willing so. Descartes notes that the passions (fear, anger, lust, etc.) are sometimes in conflict with the agent’s endeavors. The passions (as some of the body’s processes) sometimes ‘lead’ the body to perform unwanted actions. The agent, however, can detect the unwanted passion, ‘not give consent’ to it, and withhold the upcoming movement with her willpower by stopping the movement of the respective limb (ibid. art. 46).

Libet’s account on action in general resembles Descartes’ account on action to a large extent, and Libet’s interpretation of the vetoing will, more specifically, resembles Descartes’ idea of the aborting or withholding will. However, whilst Descartes describes the person (the agent) as constituted of two distinct and independent substances – the material body and the immaterial soul –, Libet does not explicitly distinguish between two distinct and independent substances. But when it comes to action, Libet indeed conceives of the agent as her own system alienated from her body but with the ability to interfere with the brain’s processes through an immaterial act of will.

---

<sup>1</sup> I am thankful to two anonymous reviewers who pointed out the ambiguity of the terms ‘interactionism’ and ‘interactivism’. In the modern cognitive sciences and in the philosophy of embodiment, these terms are used differently. Bickhard (2009), for instance, understands interactivism as a theoretical system, according to which mental phenomena, such as perceptions, emerge from the interactions of the agent with her environment. Enactivists, such as Varela, Thompson, and Rosch (2016), argue similarly and claim that cognition is essentially embodied and consists in the enactment of a world through structural coupling, that is, through the constant interaction between the agent’s body and her environment. In none of these works, interaction is understood as mind-brain, soul-body, or person-brain interaction but as the interaction between the person’s body and her environment. In this paper, however, I will use the terms ‘interactionism’ exclusively for the theory that agents act in virtue of their mind (or soul) interacting with their own body, or by interacting with their own brain (or body) directly.

This kind of dualist brain-person-interactionism is implicit in many parts of Libet's work. To begin with, Libet assumes that whatever originated in the brain, cannot have been originated in the agent. In a reply to Libet's work, Velmans argues that Libet's veto could still be a free choice *of the agent* even if it has been initiated unconsciously in the brain and just became conscious (Velmans 1991, 705). Libet disagrees with him and remarks that, in this case, the agent would not be in *control* of the action. She would only be *aware* of an originally cerebrally initiated choice (Libet. 1999, 52). Velmans, similar to MacKay (idem. 1985, 546), wants to defend a position according to which the agent is *embodied* in the processes of his own brain so that it does not matter if her choice was caused by unconscious neural processes or not. Simply put, the agent *is* nothing but her body. Libet, however, does not even consider this option. For Libet, brain and agent are two mutually independent systems, with each system having its own independently unfolding processes. The agent is primarily just an observer of all of her brain's neural activities (including her urges, desires, choices, intentions, etc.).

In spite of this observational stance, Libet grants the agent indeed some 'control of her (bodily) action', as he calls it. However, for Libet, being in 'control of the action' does not mean that the agent *takes over* some of the body's processes and *does* them.

There are many ways in which the word 'control' can be used. The most common way, however, implies some kind of interaction. For *A* to control *B* in an interactionist sense requires that *A* and *B* are distinct systems so that the processes happening in *A* and *B* take place independently from each other. However, *A* can control *B* by performing a 'counter-action' that interferes with *B*'s movement so that *B*'s movement stops or becomes aligned in the way *A* wants. If *A* does not perform such a counter-action to control *B*, *B* would just continue moving. I assume that people often use the word 'control' in this way to describe the control people exert on objects such as vehicles or animals. An agent can, for instance, control her car by hitting the breaks or by operating the stirring wheel (counter-action) to change the direction of the car's movement.

Libet uses the word 'control' in this sense to describe an interfering and regulating kind of control. For him, 'being in control of her actions' means that the agent can control the upcoming volitional process that was initiated by unconscious cerebral processes by vetoing them. This becomes clear in various formulations (idem. 1983, 641 and 1985, 538-539, and 1999, 51). The veto is the 'counter-action' that the agent 'performs' to stop or abort the progress of the volitional process. Again, none of the things that happen in and with the body including the decision and the action itself is ever *done* by the agent. The body (including the brain) is a self-contained system that develops decisions and acts in accordance with these decisions. If the agent never exerted a veto (because she is happy with the unconsciously made decisions), the body would just continue moving and acting on its own. The veto, in turn, is not an act of the body but an act of the agent alone interfering with neural processes. Or in other words, the agent is 'in' her body like a driver in her car, according to Libet, and the only thing that she can do to determine the movement of the car is hitting the breaks.

The dualist nature of Libet's person-body interactionism becomes apparent in the passages when he discusses the nature of the veto. Interestingly, Libet leaves it open

whether his account of the vetoing will be materialist or dualist (idem. 1985, 538). However, his criticism against Velmans shows that the veto cannot have been initiated unconsciously. It must be an act that “[...] can appear without prior initiation by unconscious cerebral processes [...]” (ibid., also see idem. 1999, 51-52). According to Wood and Roskies, a conscious act that has no neural signature and is not mediated by any physical process, as Libet describes it, can also not be embedded in the stream of material processes (idem. 1985, 557; Roskies 2011, 15). Instead, it must be a purely immaterial act unfolding parallelly to the brain’s neural processes. Now, if, for Libet, the agent is just an observer of her own cerebral activities, and if the only means by which she can influence these activities (the veto) is immaterial, we can assume that Libet conceives of the agent herself as an immaterial entity attached to her own body.

### 3 The Veto in the Study of Schultze-Kraft et al.

Many have rejected the various dualist elements in Libet’s works. Schultze-Kraft et al., for instance, insist that the veto, whatever it is, must be mediated by a cortical region and therefore be embedded in the other (material) activities of the brain (Schultze-Kraft et al. 2016, 1084). They also refer to a famous older experiment conducted by M. Brass and P. Haggard in which such a determination was done (Brass and Haggard 2007).

Researchers such as MacKay (Libet 1985, 546) and Nelson (ibid., 550) also criticize the interactionist aspect of Libet’s veto according to which acting consists in an aborting act *of the agent* against some of her own cerebral processes. According to them, this interactionism is not consistent with the phenomenological character of acting. Nelson notes that “[...] conscious voluntary control is part of a conscious stream parallel to, but not interacting with, cerebral processes” (ibid., 550). MacKay agrees that decisions have various neural correlates, and that agents can determine the course of their action by controlling it. However, it is a category mistake to assume that they control their action by interacting with these correlates (ibid., 546). Schultze-Kraft et al., in turn, do not reject the interactionist paradigm in their paper. In fact, they use an analogy to describe the function of the veto that reminds us of a clear person-body interaction:

“One important question is whether a person can still exert a veto by inhibiting the movement after onset of the RP [...] The onset of the RP in this case would be akin to tipping the first stone in a row of dominoes. If there is no chance of intervening, the dominoes will gradually fall one-by-one until the last one is reached. This has been coined a ballistic stage of processing. A different possibility is that participants can still terminate the process, akin to taking out a domino at some later stage in the chain and thus preventing the process from completing” (Schultze-Kraft et al. 2016, 1080).

If we take the analogy seriously, the neural activities caused by the RP and culminating in the action (the row of dominoes) are processes that unfold independently from the agent herself (the domino master). The agent, however, might have a chance to stop this unfolding process by vetoing it (take out a domino). Otherwise, the process unfolds, and the body just starts to move (the dominoes continue falling).

But even if we do not take the analogy seriously, the interactionist paradigm underlies many formulations of the study paper, especially the verbs, that Schultze-Kraft et al. use throughout the paper to describe the function of the veto. Before I begin with a linguistic analysis of their formulations, let me describe the setup of the experiment.

Schultze-Kraft et al. conducted a stop-signal experiment that they describe as a ‘duel’ against a computer (ibid.). Unlike the test subjects in Libet’s studies, the participants in this experiment were not asked to abort their bodily movement spontaneously but as a response to a computer signal. They were placed in front of a green button and were allowed to press it with their foot whenever they felt like, as long as the button is still green. However, if the button turned red suddenly, they were no longer allowed to move. A computer tried to ‘predict’ the time of the test subjects’ button press based on the collected EEG data of the test subjects. If it predicted an upcoming button press, the button turned red. In the case of an *early cancellation*, the test subjects intended to press the button; and after the button suddenly turned red, they managed to not move at all. In the case of a *late cancellation*, in turn, the test subjects intended to press the button, too; however, after the button turned red, they were not able to not move. They moved a bit. Some test subjects even touched the button. This latter scenario was called ‘*completed button press*’. According to the results of the study, there is a ‘point of no return’, after which agents are unable to not move, namely 200 ms before movement onset. However, that also means that, as the authors of the study note, “[...] a decision to move can be cancelled up until 200 ms before movement onset” (ibid., 1084).

Note that I chose neutral verb phrases to describe the target-phenomenon of the stop-signal experiment, namely ‘being able not to move’ and ‘not being able not to move’. The formulations that the authors of the study use to describe the function of the veto resemble the formulations that Libet also used: ‘Participants can/cannot *prevent* themselves *from moving*’, ‘participants can/cannot *veto/cancel/abort* prepared movement, upcoming movement, overt signs of movement, movement plans’, ‘test subjects also *exerted a veto*’, and ‘the stop process can/cannot catch up with the go process’. For Schultze-Kraft et al., all these acts of veto indeed have underlying neural mechanisms. However, the veto is still described as a (successful or unsuccessful) act *of the agent* directed *against* her own intentions and her upcoming or unfolding movements.

Consider the ditransitive verb ‘prevent (from)’ in “[...] an explanation is needed to clarify why people cannot prevent themselves from moving [...]” (Ibid., 1083). *Prima facie*, using the verb implies a distinction between two things – the one who prevents and the one who is prevented from doing something. The police can prevent the thief from running away by handcuffing him. I can also prevent myself from doing something, for instance, from sneezing loudly by squeezing my nose and mouth shut. In this case of holding my sneeze, I, as the agent, act on my own body. In both cases of prevention, there are two distinct and antagonistic processes involved. The counter-action (handcuffing, squeezing the nose), that the one who prevents performs, is a process distinct from the process that is supposed to be prevented (running away, sneezing loudly). If no counter-action would have been performed to prevent the other process, the other process would just have continued unfolding. This shows that acts of prevention fall in the category of interactionist control, as I described it in the previous chapter.

But is this plausible? Did the test subjects prevent themselves from moving in the experiment, in a similar way in which I prevent myself from sneezing loudly?

Researchers have already determined what happens in the case of a successful action inhibition in stop-signal experiments on the sub-personal level. Logan et al., for instance, describe the phenomenon as a ‘race between a (sub-personal) go-process and a (sub-personal) stop-process’ (Logan et al. 2014, 68). When an agent is about to perform an action (intended or out of habit), brain signals are sent to the peripheral muscles (go-process). The subsequent perception of the stop signal (the red button), however, causes another parallel brain signal (stop-process). If the stop-process brain signals can ‘catch up’ with the go-process brain signals, they override the growth of activation of the go-process brain signals and *prevent* them from reaching a certain threshold. However, if the go-process brain signals reach that threshold earlier, the go-process brain signals ‘escape overriding’ and the limb of the person starts to move (ibid., 70).

If this model is correct, whenever the test subjects in Schultze-Kraft’s experiment managed to not move, one group of brain signals (the stop-process) overrode another group of brain signals (the go-process) and *prevented* them from culminating in overt movement. If, however, the test subjects failed not to move, the go-process brain signals were ‘too fast’ culminating in overt movement and ‘escaping’ the stop-process brain signals. So, on the sup-personal level, there were indeed two distinct and antagonistic neural processes involved in the experiment of Schultze-Kraft et al., namely the stop-process brain signals and go-process brain signals *interfering* with each other.

However, the interference between the stop-process and the go-process was an interference between two processes *only* happening on the sub-personal level. The go-process brain signals and the stop-process brain signals accompanied or underlay the action inhibition of the participants. That does not mean that *they* were the agent’s action inhibition. It would be a mistake to conclude from these research results that the participants themselves overrode the go-process brain signals (accompanying their original action intention). The stop-process brain signals (accompanying the participants’ action inhibition) overrode them. This remark resembles the criticism that MacKay and Nelson brought forth against Libet’s interaction. Like them, I do not deny that the participants in the experiment inhibited their actions, and I also do not deny that their action inhibition was accompanied by various interfering neural processes, but I deny that the participants themselves interacted with their own neural processes.<sup>2</sup>

---

<sup>2</sup> I am not the first person referring to Schultze-Kraft’s study. However, most researchers, such as Lavazza (2016) and Uithol and Schurger (2016) agree with the theoretical framework of the study and with the study results. Some researchers criticized the study but primarily for the experimental setup. Deecke and Soekader (2016), for instance remark that the action of the test subjects was not self-initiated in a proper sense because their movement intention was influenced by the presence of perceptual cues (green and red light) and also by the set time window. In the case of a real self-initiated action, the ‘point of no return’ might, therefore, be different. However, none of the researchers that I found questions, or at least discusses the general assumption that agents can *interact* with their own neural processes (in order to inhibit a movement). Many of them even adopt the interactionist vocabulary of Schultze-Kraft et al.

## 4 An Alternative Interpretation

Simply rejecting the interactionist interpretation suggested by Schultze-Kraft et al. is not enough. I also want to suggest an alternative non-interactionist reinterpretation of what the participants did when they inhibited their action.

To begin with, (failed) action inhibition is a common phenomenon. Assume that you want to throw a banana peel in your organic waste trash can. You press the pedal, the lid opens, and suddenly you realize that you forgot to put a new plastic bag in the trash can. You immediately stop moving. However, it is possible, since you were already in the act of throwing, that your hand still releases the peel accidentally. How would you describe what happened? Certainly, in the case of a successful action inhibition, you, upon realizing that there is no plastic bag in the trash can, simply changed your action intention from throwing the peel to holding still, and you successfully held still. You slowed down your current movement and finally stopped it. In the case of an unsuccessful action inhibition, you also changed your action intention from throwing the peel to holding still. However, your hand still released the peel. That is so, because the go-process brain signals, that started in the moment when you opened the lid, could not be overridden by the stop-process brain signals, that started in the moment when you saw the empty trash can, and already reached a threshold causing your muscles to move.

I assume that the test subjects in the Schultze-Kraft experiment experienced something similar. In the case of an early cancellation, they first intended to press the button; they almost began to move; but, upon seeing the button turning red, their intention to move became an intention to hold still, and they held still. There is no interaction implied in this interpretation. The test subjects simply replaced one intention early enough with another intention resulting in a non-movement of the foot. In the case of a late cancellation, the test subjects also intended to press the button, and, upon seeing the button turning red, their intention to move became an intention to hold still. However, the brain signals accompanying the new intention to hold still could not 'catch up' with the brain signals accompanying the intention to move. The latter brain signals reached a certain threshold and thereby caused a short and sudden twitch in the foot.

This twitch might have given the test subjects the feeling of being alienated from their body. But that does not mean that body-person interactionism is correct, and that the test subjects tried but failed to veto or abort a neural process starting in their brain and culminating in the respective foot movement. It rather means that, as Velmans and MacKay argued, agents are *normally* embodied in their own body with all its neural processes. However, this feeling of embodiment can also be interrupted if neural processes accompanying an earlier action intention cause a muscle contraction that does not conform with the agent's current action intention. This non-conformity, in turn, is grounded in the delay of those neural processes that accompany the agent's earlier action intention and the slowness of her current intention's neural processes. And this is exactly what happened to the test subjects in the case of a late cancellation.



## 5 Conclusion

The purpose of this paper was to reveal the interactionist paradigm in a study conducted by Schultze-Kraft and to provide a non-interactionist reinterpretation of the test subjects' action inhibition. I did so by showing that we should not confuse what the test subjects did when they inhibited their action with the neural processes that accompany their action inhibition. Agents do not interact with their own body (or brain) when they inhibit an action. They rather change their intention or the course of their action. Unless the inhibition fails, they are fully embodied in their own moving limbs.

## References

1. Bickhard, M. H.: The interactivist model. *Synthese* **166**(3), 547–591 (2009).
2. Brass, M.; Haggard, P.: To do or not to do: the neural signature of self-control. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **27**(34), 9141–9145 (2007).
3. Deecke, L.; Soekadar, S. R. (2016): Beyond the point of no return: Last-minute changes in human motor performance. *Proceedings of the National Academy of Sciences of the United States of America* **113**(21), E2876.
4. Descartes, R.: *Die Passionen der Seele*. (French) [Les Passions de l'âme]. Meiner: Hamburg [1649] (2014).
5. Kornhuber, H. H.; Deecke, L.: Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv für Gesamte Physiologie* **284**, 1-17 (1965).
6. Lavazza, A. (2016): Free Will and Neuroscience: From Explaining Freedom Away to New Ways of Operationalizing and Measuring It. *Frontiers in human neuroscience* **10**, 262.
7. Libet, B.: Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, **8**(4), 529-539 (1985).
8. Libet, B.: Do We Have Free Will? *Journal of Consciousness Studies* **6**(8-9), 47-57 (1999).
9. Libet, B.; Gleason, C. A.; Wright, E. W.; Pearl, D. K.: Time of conscious intention to act in relation to onset of cerebral activities (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain* **106**, 623-642 (1983).
10. Logan, G.D.; Van Zandt T.; Verbruggen F.; Wagenmakers E.J.: On the Ability to Inhibit Thought and Action: General and Special Theories of an Act of Control. *Psychol Rev.* **121**(1), 66-95 (2014).
11. Roskies, A.: Why Libet's Studies Don't Pose a Threat to Free Will. In: *Conscious Will and Responsibility. A Tribute to Benjamin Libet*. W. Sinnott-Armstrong and L. Nadel (eds.). Oxford: University Press (2011).
12. Schultze-Kraft, M.; Birman, D.; Rusconi, M.; Allefeld, C.; Görgen, K.; Dähne, S.; Blankertz, B.; Haynes, J.-D.: The point of no return in vetoing self-initiated movements. *Proc Natl Acad Sci USA* **113**(4), 1080–1085 (2016).
13. Uithol, S.; Schurger, A. (2016): Reckoning the moment of reckoning in spontaneous voluntary movement. *Proceedings of the National Academy of Sciences of the United States of America* **113**(4), 817–819.
14. Varela, F.; Thompson, E.; Rosch, E.: *The Embodied Mind Cognitive Science and Human Experience*. Cambridge: MIT (2016).
15. Velmans, M.: Is human information processing conscious? *Behavioral and Brain Sciences*, **14**(4), 651-669 (1991).