

Brain Determinism and Free Will

Sergio Filippo Magni
Università di Pavia
Dipartimento di Filosofia
filippo.magni@unipv.it

ABSTRACT

The article focuses on the issue of brain determinism, by examining two recent neuroscientific experiments (Libet's and Haynes' experiments about conscious acts). Such experiments aim to show a cerebral determination of the agent's free choices. The author argues that even if their conclusions were true, that would not eliminate the common use of the concept of free will and the attribution of moral responsibility.

0. Introduction

In the marvellous philosophical *Epilogue* of *War and Peace*, Lev Tolstoy gives the following account of the issue of free will: "The problem is that regarding man as a subject of observation from whatever point of view - theological, historical, ethical, or philosophical - we find a general law of necessity to which he (like all that exists) is subject. But regarding him from within ourselves as what we are conscious of, we feel ourselves to be free. [...] However often experiment and reasoning may show a man that under the same conditions and with the same character he will do the same thing as before, yet when under the same conditions and with the same character he approaches for the thousandth time the action that always ends in the same way, he feels as certainly convinced as before the experiment that he can act as he pleases."

Nowadays the issue of free is still debated along similar lines: on the one hand we have the scientific investigation, which assumes that there is a deterministic relationship between events; on the other hand we have the manifest image, our feeling and thinking of ourselves as free to choose otherwise.

However, this contrast is still not sufficient if we wish to highlight a real problem. Borrowing another image from Tolstoy's *Epilogue*, there is still a tension between the geocentric astronomical theory, which is the manifest image, the common sense image, and the Copernican theory; yet this tension is not perceived by anyone as a difficulty. The problem in relation to free will arises because, unlike for astronomy, there are not scientific or philosophical arguments that conclusively solve the question. That is to say, the arguments that are brought forward as solutions of the problem remain speculative, mere theoretical hypotheses, not susceptible to empirical confirmation, nor immune to

philosophical counter-arguments. Yet, according to some scholars, the above would no longer be true, as recent neuroscientific discoveries have led us to the conclusion that even the choice between determinism and free will is like the choice between Ptolemaic and Copernican astronomy.

1. *Determinism and free will*

As a first step it is useful to define the main concepts involved. Generally speaking, determinism claims that each event is determined by the occurrence of sufficient conditions for its happening. Depending on how these sufficient conditions are interpreted, determinism can take several forms: theological determinism, logical determinism, causal determinism, and so on. “Determination is thus a kind of conditional necessity that can be described in various ways”¹.

In a broad sense, therefore, determinism claims that: an event B is determined if and only if there is an event A, such that necessarily if A occurs, then B occurs. In modal logic: $\forall B, \exists A \text{ Nec } (A \rightarrow B)$.

Roughly, causal determinism maintains that an event is determined by a cause. Usually, the concept of event is interpreted as including external phenomena and movements of the body as well as thoughts and mental states, and the concept of cause is interpreted as a sufficient condition that happens before the occurrence of the event² In this way, whenever the cause occurs, the effect occurs (without excluding, however, that it may also occur under other conditions: it is a sufficient condition, but not a necessary one). Moreover, a cause can be an event (or a sequence of events) of different types: physical, chemical, biological, psychological, and so on; thus we can speak of physical determinism, chemical determinism, biological determinism, psychological determinism, and so on.

In order to avoid the difficult notion of a cause, Peter Van Inwagen prefers another formulation of determinism, with a different terminology, and replaces the concepts of “cause”, “effect” and “event” with the concepts of “state of the world”, “instant of time” and “laws of nature”³. According to this formulation, determinism claims that: a state of the world, B, at time t1 necessarily follows from the state of the world, A, at time t0, in connection with the laws of nature. In modal logic: $\text{Nec } ((A \wedge L) \rightarrow B)$.

However, it is not only the concept of determinism that needs clarification; the concept of free will also requires further analysis. It would be a mistake to think that it consists only in the opposition to determinism, in the mere identi-

¹ Kane 2002, p. 6.

² For an useful analysis of the notion of cause, see Honderich 1988, cap. 1.

³ See Van Inwagen 1983, pp. 58 ff.

fication with indeterminism, with the lack of determination of the will. Lack of determination is not the same as freedom, but as chance, as Hume already said. Even if indeterminism were to be confirmed by scientific evidence, that would not mean that we could automatically infer the existence of free will: an uncaused act is not a free act, but an accidental act, completely unrelated to the agent's previous actions and choices, to his or her character and personality. Therefore, the terms involved are threefold: determination, chance and freedom.

This gives rise to a problem of intelligibility of the notion of free will in relation to that of chance, which has led to the identification of at least two features, only the first of which is implied in the notion of chance⁴:

a) the existence of alternate possibilities: the possibility to want or choose otherwise, differently from the way we actually want or choose. The agent who wants in just one possible way and has no other options available is not a free agent.

b) The agent's control over choice. That is to say, the choice has to happen in the subject (as the origin and ultimate source of it) and be under his control, because it is determined by internal factors (endogenous), and not by external conditions (exogenous) that the agent cannot control. Such a condition is not observed by chance. The will must then be in our power, up to us. As someone has pointed out, the problem of free will is "the issue of how much control over our own actions, and so over our own lives, the extrapolitical world - the natural (and perhaps also the supernatural) world – affords us"⁵.

2. Brain determinism: Libet's and Haynes' experiments

In recent years, many scholars have emphasized that the sciences studying the functioning of the human brain (neurophysics, neurochemistry, neurobiology, etc.) would lead us to a deterministic view. Scientific analysis of the brain shows that the behavior of brain cells, neurons, is completely determined: each neuron receives electrochemical inputs from its roots, called 'dendrites', and relays outputs of the same nature, through a trunk, called 'axon', which ends in connection with the roots of other neurons, called 'synapses'. The synapses stimulate or inhibit the activity of other neurons through the release of chemical elements, called 'neurotransmitters', which control the passage of electrical impulses to other synapses. Each thought or mental activity involves millions of neurons, but this would be completely determined by physical and chemical laws⁶.

⁴ See Watson 1987, p.145.

⁵ Ekstrom 2000, p. 3.

⁶ See Honderich 1993.

The subjective awareness of free will, the inner experience of freedom that we have as agents would therefore be misleading, and not supported by reliable scientific data. It would be nothing more than an illusion, although it plays an important role in common daily interactions, and can itself be explained by physical and natural parameters. Something similar, as it was said, to the geocentric astronomical view.

Such a theory was often held over the centuries by determinist philosophers. What is new today is that this theory seems to be confirmed by some relevant neuroscientific experiments, that, according to someone, would definitively solve the problem. Two experiments, in particular.

The first one is an experiment on voluntary actions (spontaneous actions that the subject feels under his own control) proposed by Benjamin Libet. The experiment shows that these kinds of actions are accompanied, in the cerebral cortex, by an electrical activity that precedes the time when the subject becomes aware of his or her own decisions⁷. Electrical brain activity of the cerebral cortex is recorded while subjects are asked to move a hand as soon as they feel the urge to do so, remembering the time of conscious decision by observing the dial of a special clock 25 times faster than a normal clock. This action is preceded, by 1000 to 550 milliseconds, by a negative brain potential (the so called ‘readiness potential’), that originates from the brain area involved in motor preparation: supplementary motor area (SMA). The subject, however, becomes aware of his or her own decision only about 400 milliseconds before the act, and therefore only after the onset of readiness potential in that area.

The experiment would show that the brain has taken its decision before the person becomes aware of it. As Libet writes: “the initiation of the freely voluntary act appears to begin in the brain unconsciously, well before the person consciously knows he wants to act!”⁸

However, Libet’s experiment does not solve any problems. First of all, the readiness potential is generated by ‘supplementary motor area’ (SMA), but more fundamental planning stages might be involved in other areas of the cerebral cortex. Moreover, the time delay between the onset of readiness potential and the subject’s awareness is very short, and it could lead to measurement inaccuracies. Finally, since in the experiment there is no choice between different options of action (but only between moving or not moving the hand), the readiness potential might be a merely unspecific preparatory activation, and not be related to the outcome of the choice. Instead, we would consider it related to the that outcome if it enabled us to foresee it.

⁷ See Libet-Gleason-Wright-Pearl 1983. For a general account of the neuroscientific analysis of free will, see Libet-Freeman-Sutherland 1999.

⁸ Libet 2002, p. 555.

In order to solve these problems John-Dylan Haynes and others have recently proposed a more accurate test, which would validate the results of Libet's experiment. Whereas a letter stream updated every half second is shown to the subjects, they are asked to consciously choose whether to press a left or a right button by using the corresponding index fingers. Subjects are asked to indicate the time of their motor decision, while remembering which letter was presented when they consciously took their motor decision. Sophisticated neuro-imaging techniques (functional magnetic resonance imaging, fMRI) enables us to measure brain activity of various brain regions and to locate it with good precision.

The experiment shows that there are two other brain's regions that present electric activity before the agent's conscious choice, and the observation of which zone of each of the two regions is involved enables us to foresee which choice will be made. The first is a part of the frontal cortex (the frontopolar cortex), in which there is a predictive information preceding the conscious motor decision by up to 10 seconds, the second is a part of the parietal cortex (an area between the precuneus and the posterior cingulate cortex) that shows predictive activity after the activation of the frontal cortex. The supplementary motor area, SMA, is involved only later.

According to its authors, the experiment "suggests that when the subject's decision reached awareness it had been influenced by unconscious brain activity for up to 10 s" and provides "a tentative causal model of information flow, where the earliest unconscious precursors of the motor decision originated in frontopolar cortex, from where they influenced the buildup of decision-related information in the precuneus and later in SMA, where it remained unconscious for up to a few seconds"⁹.

Haynes' experiment seems to overcome the weaknesses of Libet's experiment. It allows to identify high-levels motor control regions, which are more fundamental than SMA; to record a much longer time between brain activity and awareness of decision, so reducing timing inaccuracies in reporting the onset of awareness; and, above all, it allows to foresee what will be the outcome of the subject's actual choice, that appears to be determined by specific areas of the prefrontal cortex.

According to both experiments, therefore, subjective awareness arises only after an unconscious brain decision process has started, and that would exclude the role of conscience in choosing. The choice seems to be determined by electro-chemical factors outside the intentional control of the subject. So, the agent would not have any control over his choices and these would not have alternate possibilities (the conditions of human free will, as we have seen).

⁹ Soon-Brass-Hainze-Haynes 2008, p. 3.

3. *Brain determination and compatibilism*

A debate about the interpretation of the two experiments has been going on for several years now. Both experiments admit interpretations that would eventually save the notion of free will. It might be argued, for example, that the experimenter excessively determines the type of choice available to the agents; that the type of choices required by the experimenters is not analogous to common choices, intrinsically related to preferences and reasons; that even the agents' usual choices are not always conscious; that the cerebral causal connections are very complex, and so on. Libet himself has brought forward an interpretation of his experiment that does not completely eliminate free will, for it gives to the agent the power to stop the execution of the choice: the choice would be determined by cerebral activity, but the agent could stop the execution of it. There would be no free will, therefore, but something as a free veto: "the conscious veto may *not* require or be the direct result of preceding unconscious processes. The conscious veto is a *control* function different from simply becoming aware of the wish to act"¹⁰ (assuming, implausibly, that the veto awareness should not be preceded by any readiness potential).

What should be stressed, however, is that, even if we admit that those experiments actually show a cerebral determination of choice and exclude the role of subjective consciousness in the decision, they cannot exclude free will, and the related concepts of control and alternate possibilities. Analogously, they do not result in a disregard of the difference between situations where this freedom actually appears to exist and situations where it does not appear to exist. In other words, they do not invalidate the claim that cerebral determination is compatible with free will. At least they do not invalidate it in a weaker sense of the term 'free will', which is sufficient, however, to preserve the meaning of the expressions commonly associated with it: "I can choose freely", "I could have chosen otherwise", and so on.

If we keep this weaker sense of the term, free will means the possession of two requirements existing even in the presence of a cerebral determination of choice. The first requirement is the absence of external coercion to choose, namely the *opportunity* to choose according to one's own preferences (along the lines of the more widespread and ancient form of compatibilism, dating at least to Hume). The second requirement is the presence of internal psychophysical characteristics, i.e. the *capacity* to choose according to one's own preferences (along the line of a more recent form of compatibilism)¹¹.

¹⁰ Libet 2002, p. 558.

¹¹ See Kenny 1975 e 1989. For an account of this type of compatibilism, see Magni 2005, cap. IV.

Even if the agent's choices were completely determined by brain mechanisms, the difference between agents who have the opportunity and the capacity to choose otherwise and agents who do not have that opportunity and capacity would still remain. Even if the agent's choices were causally determined, it would still be true that, at a given moment and in relation to a particular choice, he could have chosen otherwise, having both the opportunity and the capacity to choose, but, at another time or in relation to another choice, he could not have chosen otherwise because, for example, he was physically or mentally disturbed. Even in a deterministic universe, it does remain true that I could have chosen otherwise, because I was not forced by anyone and in full possession of my own mental faculties, while somebody else could not, because he was coerced or mentally disabled.

This version of compatibilism succeeds in explaining the two conditions of the notion of free will that we have considered above: the agent's control of choice and the presence of alternate possibilities.

The control can be identified with the agent's capacity of choice: an agent who has no control over his choice is also unable to choose, and conversely those who have no capacity to choose have no control over their own choice. As it was noted by Dennett, to have control over something implies being able to exercise that control. Such a capacity requires a certain amount of knowledge, so that the degree of knowledge generally reflects the degree of control: the more complete is the knowledge, the more complete is the control. And it requires a certain capacity to correct one's behavior, i.e. a capacity of self-control, which enables the agent to review his or her projects and goals in order to obtain a better outcome. But such conditions are compatible with determinism, just as the control of a model aircraft through a radio control is compatible with determinism¹².

Accounting for the second condition - alternate possibilities - might seem more complex. The Consequence Argument used by Van Inwagen to criticize compatibilism would show how compatibilism fails to account for this condition. If determinism is true, according to Van Inwagen, the present state of the world necessarily follows from the past state of the world, in connection with the laws of nature. An alternate possibility, that is a different state of the world, could exist only if the laws of nature or the past states of world could change; yet, the agent does not have the power to change the laws of nature, which are inescapable, nor to change the past states of the world, which are unalterable. The agent could never act and choose otherwise than he has actually acted or chosen. "If determinism is true – Van Inwagen writes -, then our acts are the consequences of the laws of nature and events in the remote past.

¹² See Dennett 1984, pp. 52 ff e 81 ff.

But it is not up to us what went on before we were born, and neither is up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us”¹³.

Yet, the word ‘power’ is an ambiguous word. The Consequence Argument seems to be irrefutable, if we use ‘power’ in Van Inwagen’s sense: ‘possible’ is what is not inconsistent with the past and the laws of nature (something like physical modality). But ‘power’ can also be used to mean other modalities: for example the so-called ‘dynamic modality’ (opportunity and ability)¹⁴. It is true, for example, that given the past and the laws of nature, this afternoon I *cannot* take a walk and I am determined to stay at home; however, even in such case, it is appropriate to say that, in another sense of the word ‘power’, I *can* take a walk because I am not coerced nor disabled, while somebody else, who is coerced and disabled, *cannot* take a walk, although both of us are determined by the past and the laws of nature.

The physical modalities of possibility and the dynamic modalities of possibility are indeed different, as it is shown by the fact that they have a different logical treatment. Modal logic of capacity does not have two axioms of modal logic of possibility. The first is the so-called ‘axiom of possibility’ (*ab esse ad posse valet consequentia*): if something happens, then it must be possible (in modal logic: $A \rightarrow PosA$). This principle does not hold in relation to capacity: we can hit a target at the first shot, but we can do so accidentally and therefore without having the capacity. The second is the so-called ‘axiom of distribution of possibilities over the disjunction’: it states that if A or B are possible, then A is possible or B is possible (in modal logic: $Pos(A \vee B) \rightarrow (PosA \vee PosB)$), and even this principle does not apply to capacity. The fact that, upon request, I can take a red card or a black card from the deck does not mean that I can take a red card upon request or a black card upon request.

Thus, the modality of the verb ‘power’ which is implied in talking about physical impossibility is not the same as that implied in talking about human capacity. Even if there exists a strong determinism and the past and the laws of nature actually determine our choices, we can talk, in that sense, about alternate possibilities.

Holding this view does not mean ignoring the results and methods of scientific knowledge. This kind of compatibilism can indeed be understood in a naturalistic way, which explains the agent’s capacity to choose otherwise as a result of the evolutionary history of mankind: a trait which was selected in the evolution of animal species because it proved to be essential to the process of the human animal’s adaptation to its environment¹⁵. Furthermore, it allows us to speak of degrees of free will and of the possession of a greater or lesser free-

¹³ Van Inwagen 1983, p. 16.

¹⁴ According to Kenny’s terminology, see Kenny 1978, p. 131.

¹⁵ See Dennett 2003.

dom to choose, depending on the degree of possession of this capacity of choice. Similarly, this kind of compatibilism allows us to explain the acquisition or the loss of free will, depending on the increase or decrease of this capacity, and explains the fact that some capacities may be more fundamental than others, as they are the conditions of the possession of more superficial capacities. It opens, namely, to the consideration of different levels of capacity, according to the dependence of a capacity on the previous possession of other capacities.

Yet, this kind of compatibilism can do all these things without being tied to determinism. Whether determinism is true or false is an empirical question on which this type of compatibilism remains agnostic, at least until an unquestionable scientific evidence allows us to solve the problem.

4. Cerebral determination and responsibility

Since free will is commonly supposed to be a necessary condition for moral (or legal, political, etc.) responsibility, if cerebral determination is compatible with free will, it is also compatible with the attribution of moral responsibility to the agent. Nonetheless, the question of compatibility between determinism and freedom has to be kept separate from the question of compatibility between determinism and responsibility. That is, there are two possible compatibilisms (or incompatibilisms): of freedom and of responsibility; and the two views do not coincide. While those who hold a compatibilistic view in regard to freedom are also compatibilists in regard to responsibility, not all incompatibilists in regard to freedom are also incompatibilists in regard to responsibility¹⁶.

Therefore, the statement that if there were no freedom we could not talk about moral responsibility is not a good argument in favour of free will. This is so at least for two reasons. First, there may be philosophical counter-arguments that call this statement into question, as it is shown by a thought experiment proposed by Harry Frankfurt: the case of an occult controller. It is the case of a person who can remotely control the choices of an agent, because she is able to change the agent's will without the agent being aware. Yet, the controller intervenes to change the agent's will only if it is contrary to his own, when the two wills agree the controller does not intervene: in this case the agent, who is not coerced by anyone, spontaneously performs the action, and is deemed responsible for it. However, even if the agent was held responsible, he could not act otherwise, because if he had chosen a different action, the con-

¹⁶ This is true both for some hard-determinists, who deny free will but not the attribution of moral responsibility, both for some libertarians, who are incompatibilists about free will and compatibilists about responsibility, as in the case of semi-compatibilism of John Martin Fischer: see Fischer 2002, p. 306.

troller would have intervened and changed his will. Thus, Frankfurt concludes, we can be held responsible for any action, even when we do not have the possibility to do otherwise; so we do not have free will¹⁷. Such an argument has opened a long debate.

The second reason is a claim which has been not much underlined in the debate of recent years: there is not a single conception of moral responsibility but there are at least two, and if, taken in one sense, responsibility may be considered incompatible with determinism (provided that Frankfurt's counter-argument does not work), in the other it does not. The two conceptions are the retributionistic view of responsibility and the consequentialistic (pragmatic, utilitarianistic) view of responsibility. If what matters in the attribution of responsibility is not the compensation for a merit or a fault (as it is claimed by the retributionistic view), but the prevention and correction of behavior (as it is claimed by the consequentialistic view), it seems clear that responsibility may well coexist with determinism; and, indeed, the deterministic view is one that can better explain the effectiveness of attributing responsibility. The attribution of responsibility (of a punishment or a reward) would determine the course of future events, preventing or correcting the occurrence of certain actions. In fact, it was the conception of responsibility most often supported by determinists (from Hobbes to, nowadays, Honderich).

At this point one might object that only the retributionistic view of responsibility is appropriate and corresponding to the common conception of responsibility, and thus it is able to grasp its ultimate or "metaphysical"¹⁸ sense. But such an objection would assume that we have already solved the problem of the justification of responsibility (in favour of a retributionistic view). However, the choice of which of the two views is the best one is not a choice that can be said to be free from general normative assumptions, even though these may be often not explicit. This choice depends on the adoption of a specific evaluative point of view; it depends on the answer we give to the question of which is the right theory to be adopted, the morally correct one. As one of the founders of neuroethics recognized, "the concept of responsibility is a social construct that exists in the rules of the society and not in the neuronal structures of the brain"¹⁹.

References

D. Dennett C. (1984), *Elbow Room. The Varieties of Free Will Worth Wanting*, Oxford, Clarendon, 1984.

¹⁷ See Frankfurt 1969.

¹⁸ See Ekstrom 2000, p. 150.

¹⁹ Gazzaniga 2005. p. 99.

- D. Dennett C. (2003), *Freedom evolves*, New York: Viking, 2003.
- Ekstrom L. W. (2000), *Free Will. A Philosophical Study*, Boulder and Oxford: Westview, 2000.
- J. Fischer M., (2002), "Frankfurt-Type Examples and Semi-Compatibilism", in R. Kane ed., *The Oxford Handbook of Free Will*, Oxford, Oxford Un. AP, 2002.
- Frankfurt, H. (1969), "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy*, 66, pp. 829-839.
- Gazzaniga M (2005), *The Ethical Brain*, New York-Washington, D. C., Dana Press, 2005, tr. en. Mind ethics. Turin, Code, 2006.
- Honderich T. (1988), *A Theory of Determinism*, Oxford, Clarendon Press, 1988.
- Honderich T. (1993), *How Free are You?*, Oxford, Oxford University Press.
- R. Kane (2002), "The Contours of Contemporary Free Will Debates", in R. Kane, ed., *The Oxford Handbook of Free Will*, Oxford, Oxford Un. AP, 2002.
- Kenny A. (1975), *Will, Freedom and Power*, Oxford, Blackwell, 1975.
- Kenny A. (1978), *Free Will and Responsibility*, London: Routledge & Kegan Paul, 1978.
- Kenny A. (1989), *The Metaphysics of Mind*, Oxford, Clarendon Press, 1989.
- Libet B. (2002), "Do We Have Free Will?", In R. Kane, ed., *The Oxford Handbook of Free Will*, Oxford, Oxford Un. AP, 2002.
- Libet B. - C. Gleason - E. Wright - Pearl D. (1983), "Time of Conscious Intention to Act in Relation to Cerebral Potential", *Brain*, 106, pp. 623-42.
- Libet B. - Freeman A. - Sutherland K, eds., (1999), *The volitional Brain: Toward a Neuroscience of Free Will*, Thoeverton, Imprint Academic, 1999.
- Magni S. F. (2005), *Teorie della libertà. La discussione contemporanea*, Roma, Carocci.
- Soon C.S. - Brass M. - Heinze H-J. - Haynes J-D., (2008), "Unconscious Determinants of Free Decisions in the Human Brain", *Nature Neuroscience*, Advance Online Publication, 10.1038/nn.2112.
- Van Inwagen P. (1983), *An Essay on Free Will*, Oxford, Clarendon, 1983.
- Watson G. (1987), "Free Action and Free Will", *Mind*, 96, 1987, pp. 145-172.