
Free will and modern science

The British Academy has just published a volume of papers, 'Free Will and Modern Science', illustrating the present state of the debate about whether humans have free will. The volume's editor, Professor Richard Swinburne FBA, discusses some of the issues.

THE OUTBURST OF looting, vandalism and arson in various British cities in August 2011 produced a spectrum of attitudes towards the rioters. There were commentators who claimed that the rioters were wicked; what they did was entirely 'their fault'. They were 'morally responsible' for their conduct and 'deserved' to be punished severely. Then there were commentators who claimed that, although to some extent the rioters were 'responsible' and 'deserved' punishment, nonetheless what happened was not entirely their fault. The rioters were the children of absent fathers, had never had paid employment, lived in areas with no social or cultural facilities, and were subject to the influence of gangs. So they only 'deserved' a small punishment. And finally there were those commentators who thought that the rioters were entirely the product of their environment, and 'you can't blame them for what they did'.

Implicitly (but hardly ever explicitly) each of these groups were committing themselves to a certain theory of free will. 'Free will' is a philosophers' term, and can be defined in various ways; but I think that the most useful understanding of 'free will' for this kind of context is that someone has free will if (and only if) they are morally responsible for their intentional actions. Being 'morally responsible' for our actions means being 'morally guilty' for doing what is wrong (or perhaps only what we believe to be wrong); and 'morally meritorious' for doing what is good (or perhaps only what we believe to be good), especially if we have no obligation to do it. So the first two groups of commentators were both committed to the view that the rioters had some degree of free will – though for the second group the free will was of a limited kind – while the third group would have denied the rioters had any free will at all. And that leads us straight into the two big issues about free will over which philosophers, scientists, and theologians have agonised over the past two and a half thousand years. The first issue is what we would need to be like in body and mind in order to have the requisite sort of freedom – for example, is it necessary that our actions should not be totally predetermined by our brain states? And the second issue is what we are actually like in body and mind – what kind of freedom do we have? Two of the three main philosophical positions available today on the first issue were implicit in writings 2,500 years ago; but recent philosophical discussions have made them much

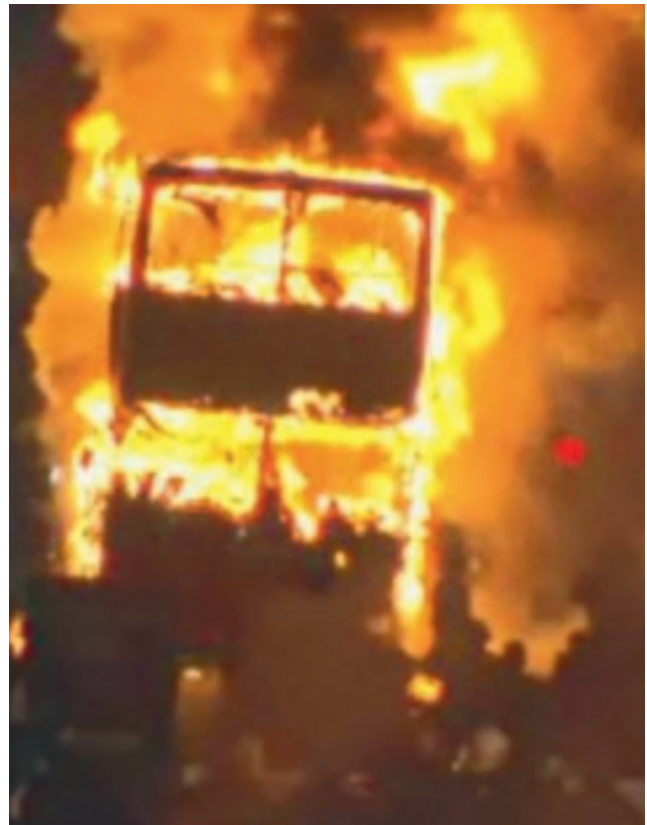


Figure 1. A double-decker bus is set ablaze, 8 August 2011. Image: VOA TV (Wikipedia Commons).

clearer and sharper and so sometimes more persuasive. But many scientists (and popular expositors of their work) think that recent scientific discoveries have put us in a far better position than any earlier generation to reach a definitive view on the second issue.

Three philosophical views

The most natural view on the first issue is to say that, in order to have the requisite kind of free will, it is necessary and sufficient that our intentional actions are not fully caused by preceding events which we do not ourselves cause. This was the view – in my opinion – of almost all Christian theologians before Augustine,¹ of Duns Scotus,

¹ See for example Irenaeus: 'God made man a free [agent] from the beginning, possessing his own power ... to obey the behests of God voluntarily, and not by compulsion of God... But if some had been made bad, and others good, these latter would not be deserving of praise for being

good, for such they were created; nor would the former be reprehensible, for thus they were made [originally]'. *Against Heresies*, 4.37, trans. A. Roberts, *The Writings of Irenaeus* vol. 2 (Edinburgh: J&T Clark, 1869), pp. 36-37.

of the Council of Trent and so of all subsequent Catholic theologians, and of all Eastern Orthodox theologians; and of course also of many atheists, agnostics and advocates of other religions. The view that the absence of full causation is a necessary condition for free will is called 'incompatibilism'; free will is incompatible with determinism. And many incompatibilists hold that the absence of such causation is sufficient to make our intentional actions free. This view nevertheless allows the possibility that our possession of free will is a matter of degree; our free will is less insofar as our actions are influenced, though not fully determined, by nature and nurture – our genes predisposing us to behave in a certain kind of way, or our environment making it difficult for us to resist peer pressure.

Then secondly there is the view, of which there are many subtle variants, that we are free if (and only if) our actions are the result of a choice which is in some way rational and not the result of 'compulsion'. So long as we are doing what we want to do and we have some reason for doing, and no one is coercing us to do the action, any action of ours is free. This is usually thought to rule out from being 'free', not merely actions that we do in response to threats (e.g., threats to kill or torture us), but also actions that we are 'psychologically compelled' to do (e.g., as a result of a drug addiction which the agent wishes that he did not have). But on this view we still have 'free will' when we do actions that we want to do and have a reason for doing, even if the effect on us of our wants and reasons is totally determined by our brain events (or by anything else such as God). This has been the view of a minority among Christians, who have thought that while God predetermines all our actions, we are nevertheless morally responsible for the bad ones. I include among this minority Augustine (in his later writings), Aquinas (in some of his writing) and some classical Protestants. But its best known philosophical exponents were Hobbes and Hume.² The view is called 'compatibilism', because it claims that free will is 'compatible' with scientific determinism.

A third view has however become prominent in recent years: the view that free will is an illusion. On this view if our actions are fully caused by previous events, we are not responsible for them; and if they are not so caused, then it is a matter of chance which actions we perform, and so again we are not responsible for them. There could not – on this view – be such a thing as moral responsibility in the stated sense, and so no one could 'deserve' to be punished. We may call this view the 'illusion view'. It could still be the case on this view that it would be good for the state to punish wrongdoers, if such punishment served a useful utilitarian function – e.g. deterring others from committing similar crimes, or reforming the criminals; but the punishment would not be 'deserved'. Each of these views on the first issue are represented in the *Free Will and Modern Science* volume.

Scientific discoveries

But what sort of free will do we have? Until recently the views which thinkers held on this issue were derived from their overall world-views. Those theologians who thought that God sent to Hell some people who knowingly did wicked actions, and that he would not do so if their actions were caused by factors outside their control, held the view that our actions were not always so caused. Those philosophers and scientists who thought that every event has a cause, drew the conclusion that all human actions are caused by a chain of events going back to events quite outside the agent's control. But in recent years, and especially in the last 25 years, two or three scientific discoveries have had a great influence on the discussions of the extent to which and the way in which our intentional actions are predetermined.

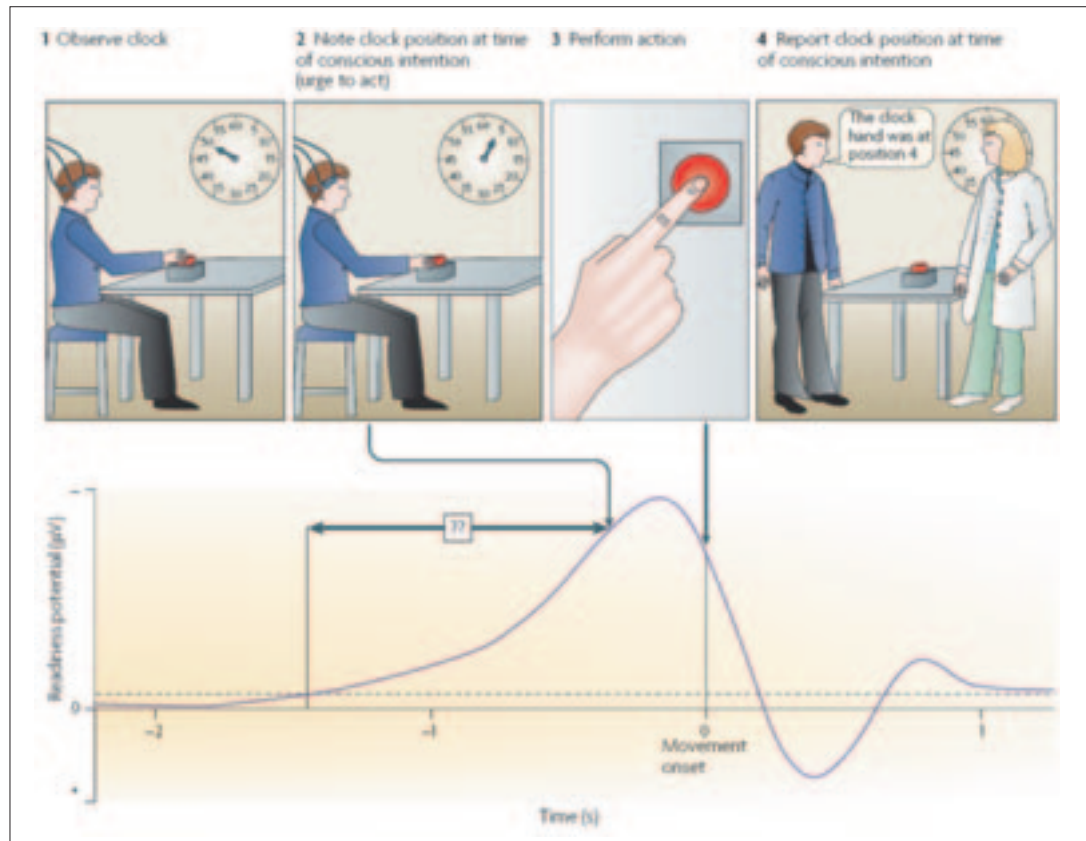
The first of these discoveries is Quantum Theory, which in its most common interpretation has the consequence that the fundamental laws of physics are not fully deterministic. Normally of course indeterminism on a small scale will average out so as to produce virtual determinism on a large scale. For example, if it were a totally indeterministic matter whether a coin landed heads or tails – if there was a physical probability (an inbuilt bias) of a half that the coin would land heads each time it was tossed, and a probability of a half that it would land tails – then in a million tosses, it would be very probable indeed that the proportion of tosses of heads would be very close to a half. But it is possible to have large-scale processes whose outcome is determined by very small-scale processes; for example scientists could construct a hydrogen bomb such that whether or not it exploded was determined by whether some atom which had a physical probability of one half of decaying within an hour, decayed within that time. Then it would be a totally chance matter (with a physical probability of a half) whether the bomb would explode. Now there is some plausibility in supposing that the brain is a system in which small-scale events not fully determined by previous brain events (and so perhaps themselves caused by uncaused decisions) cause our intentional actions. But the common interpretation of Quantum Theory remains open to dispute; and neuroscientists simply do not know nearly enough about the brain to know if the brain is a system in which small-scale brain events not caused by previous brain events cause our intentional actions. One paper in *Free Will and Modern Science* justifies this latter agnostic conclusion.

But the greatest influence on recent discussions of the second issue has come from another area of recent neuroscience. Almost everyone agrees that if we are to hold people morally responsible for their actions, it must be the case that their conscious intentions (via their brain events) cause those actions. If a rioter is to have the kind of free will which makes him morally responsible for looting a shop, it must be that he consciously intended to loot the shop; and that his conscious intention caused the brain events that caused the movements of his legs and arms that constituted

² Hobbes wrote that a person's freedom consists in his finding that he has 'no stop in doing what he has the will, desire, or inclination to do' (Thomas Hobbes *Leviathan* 2.21.) And Hume wrote that, 'when applied to voluntary actions', 'by liberty ... we can only mean a power of acting or not acting

according to the determinations of the will; that is, if we choose to remain at rest, we may; if we choose to move, we also may.' David Hume, *An Enquiry Concerning Human Understanding*, Section 8, Part I.

Figure 1. Schematic representation of the Libet experiment. Source: P. Haggard, 'Conscious intention and motor cognition', *Trends in Cognitive Sciences*, 9:6 (2005), 290-5.



looting. But a now famous series of experiments performed by Benjamin Libet in the 1980s and frequently repeated by others has been interpreted by many neuroscientists as showing that our conscious intentions make no difference to which bodily movements we make. A typical Libet-type experiment has the following form (Figure 1). Subjects sitting at desks are told to move a hand at some moment of their choice within a period of 20 seconds; during the 20 seconds they watch a very fast clock, and are told to note and subsequently report the instant at which they formed the 'intention' to move the hand. Wires attached to their skulls record changes of electric potential on the skull. Libet discovered that there was almost always a build-up of electric potential on a subject's skull half a second before the time (as reported by the subject) at which the subject formed their intention to move their hand. The very strong correlation between this readiness potential and subsequent hand movements was interpreted as showing that the brain events which caused the build-up of potential also caused the hand movements. Many neuroscientists have claimed that this result obtained from the study of quick actions in a morally unimportant situation shows that all our intentions in all situations are mere 'epiphenomena', in no way influencing our behaviour. Various new technologies for discovering which parts of the brain are active when has led to detailed experimental work in many neuroscience laboratories revealing further correlations of this kind, between prior brain events and subsequent bodily movements.

Some neuroscientists and many philosophers have, however, cast doubt on whether these experimental results show what has been claimed. We do not know whether or not even in a Libet-type experiment prior brain events producing the characteristic readiness potential sometimes occur without being followed by the bodily movements, and so whether the brain events merely indicate that the subject is considering making the movement rather than actually initiating it. And anyway even if there is a deterministic process operative, why not suppose that the brain event indicated by the readiness potential causes the intention, and the intention causes the bodily movements?

Conscious events

My own view is that a great deal depends, both for how these experiments are to be interpreted and more generally for whether our intentional actions are fully caused by brain events, on the solution to another great philosophical issue, lurking in the background. This is the issue central to the mind/body problem, of what is the nature of conscious events (sensations, thoughts, decisions etc.). If conscious events just are brain events, then everything we do – insofar as it determined – is determined merely by brain events in accord with physical laws; and that to my mind would rule out free will. But surely a visitor from another planet with a very different kind of body from ours could find out just as well as we can what is

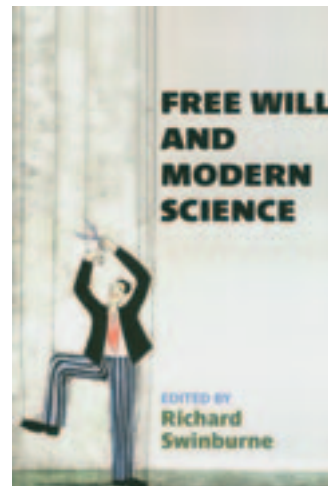
happening in our brains, but would still want to know whether we felt anything if he stuck a pin in us. What he would not know merely from the study of our brains and behaviour is what (if any) sensations, thoughts and intentions we are having; and that strongly suggests that these must be thought of as conscious events distinct from the brain events which clearly often cause them.

Given that point (constituting 'mind/brain event dualism'), it seems to me that we can only believe what the subjects in Libet-type experiments tell us about the time at which they formed their intentions (which provides the evidence for the radical interpretation of those experiments), if we believe that the subjects tell us what they do (e.g., 'I formed my intention when the clock read 4.05 secs') because they are conscious of their intentions and have the intention to tell us the truth about them. In other words, we can only justifiably come to believe that a subject's intention to move a hand doesn't cause the hand movement if we also believe that their intention to tell the truth does cause the words reporting it ('I formed my intention when the clock read 4.05 secs') to come out of their mouth. And so more generally, we can only have evidence that sometimes our intentions do not cause (via our brain events) our bodily movements, if we presuppose that sometimes our intentions do cause (via our brain events) our bodily movements. So we could never have any justification for not believing what seems to us as we act to be manifestly the case, that at least sometimes our intentions cause our brain events and thereby our bodily intentions. If these points about the separate existence of conscious events and their causal influence on the brain were accepted, it would move the discussion into the issue of how and when our brain events cause quite separate conscious events, and how and when conscious events cause the brain events which cause bodily movements. This would require a scientific theory of a totally different kind from the kind of theory normally studied by physicists, an enormous scientific revolution of a magnitude compared to which a discovery by physicists that there can be signals faster than light would be very insignificant. Hence of course the strong inclination to deny mind/brain event dualism! Just look where that takes you!

But even if our intentions cause our brain events, and some of those intentions are not themselves fully caused by brain events (which would require the brain to be an

indeterministic system), would that be enough to make us morally responsible? The mere absence of a causal chain doesn't seem to me enough to make us morally responsible. In my view it would have to be the case when our intentions are said to cause our bodily movements, the more accurate description is that we intentionally cause those movements. It would have to be the case that an agent, not a mere conscious event connected to the agent's body, does the causing, if the agent is to be morally responsible for their actions. And when we reflect on what is involved in 'trying', when we try to do a difficult bodily action – for example, to pronounce a difficult word or to lift a heavy weight – it does seem that the 'trying' just is us intentionally exercising causal influence. My view is that if and only if agents consciously cause their intentional actions and are not fully caused to do so by other events, do they have free will; and I share the view of the second group of commentators mentioned earlier that free will is a matter of degree, and that the free will of many of the rioters was of a fairly limited degree.

Richard Swinburne is Emeritus Nolloth Professor of the Philosophy of the Christian Religion at the University of Oxford, and a Fellow of the British Academy.



Free Will and Modern Science, edited by Richard Swinburne, is a British Academy Original Paperback. More information can be found via www.britac.ac.uk/pubs/
