

The Implications of Advances in Neuroscience for Freedom of the Will

Hilary Bok

Department of Philosophy and Berman Institute of Bioethics, The Johns Hopkins University, Baltimore, Maryland

Summary: Some neuroscientists argue that advances in neuroscience threaten to undermine our freedom. The argument here is that those concerns are instances of a more general concern about the compatibility of freedom with causal determinism, and that denying that our choices are fully determined

under causal laws presents a different set of problems for the claim that we have free will. An alternative account of freedom is presented, consistent with determinism in general, and with advances in neuroscience in particular. **Key Words:** Freedom, responsibility, determinism.

INTRODUCTION

When neuroscientists describe the moral implications of advances in their field, they often seem to assume that if those advances show that our choices and actions are caused by events in our brains, then it follows that we do not make those choices or perform those actions freely. Thus, Michael Gazzaniga¹ writes:

(T)here are two primary and opposing views: that we have free will, and that we don't. Those who believe in free will (indeterminists) believe that some *x* factor—whether it's the "ghost in the machine," the soul, the mind, or the spirit—allows us to make choices and determine our actions and even our destiny by acting upon and changing the physical world and our path in it. Those who don't accept free will (determinists) believe that we live in a predetermined world—whether it's caused by fate, preordination, or genetic hard-wiring—where every action, human and otherwise, is inevitable.

If Gazzaniga's description is accurate, then it is easy to see why he believes that neuroscience seems to threaten our belief in our freedom. Neuroscientists try to discover how the brain works, and in so doing they try to illuminate the physical causes of the events that occur in it. To the extent that they succeed, they leave that much less

room for any ghostly "*x* factor" to operate. If our freedom requires that our choices and actions be due to some *x* factor that allows our choices to escape causal explanation, then the more neuroscience explains, the less scope there will be for our freedom.

Even if we leave *x* factors aside, however, advances in our understanding of how the brain works can seem to threaten our conception of ourselves as free and responsible moral agents. The more we discover about the neural processes that underlie our choices, and the more we understand about their causes, the less room there seems to be for us ourselves to decide what to do, and the more our choices seem to be the inevitable results of antecedent causes. And not just our choices: as neuroscience advances, it might, for all we know, turn out to be capable of explaining our entire mental lives.

Given sufficient information about our initial state and our environment, neuroscience might eventually allow us to predict which considerations will occur to us, which ones we will take seriously and which ones we will disregard, which temptations we will be subject to and how much we will try to resist them, and what, as a result, we will choose to do. If so, then we would seem to be not the authors of our lives, but beings whose actions have simply unfolded in accordance with natural laws in ways we could not have altered. As Benjamin Libet² writes, if we are "completely defined by the deterministic nature of physical laws," then "we would be essentially sophisticated automatons, with our conscious feelings and intentions tacked on as epiphenomena with no causal power."

Address correspondence and reprint requests to: Hilary Bok, PhD, Department of Philosophy, The Johns Hopkins University, Baltimore, MD. E-mail: hbok@jhu.edu.

The basic problem is the assumption that freedom is incompatible with causal determinism: that if our choices are caused by antecedent events in accordance with some set of causal laws, then they cannot be free. Both Gazzaniga and Libet make this assumption and it is because they make it that they see causal explanations of our choices as threats to free will. This assumption is not obviously true, however, and many philosophers reject it.

The debate over whether or not we can be free if our choices and our actions are determined according to causal laws has a long history in philosophy.^{3,4} Long before scientists began to provide plausible and informative explanations of events in the brain, philosophers and scientists recognized that such explanations were possible in principle and that they raised serious questions about whether, if our choices and our conduct were fully determined according to causal laws, we might nonetheless be said to be free. From the start, many philosophers argued that causal determinism is fully compatible with freedom and moral responsibility and, although this debate continues, many philosophers accept these arguments today. In this article I want to explain why.

PROBLEMS WITH UNDETERMINED CHOICES

The assumption that freedom is incompatible with determinism states a negative condition on freedom of the will: if determinism is true, then we are not free. It does not tell us anything positive about what freedom of the will might be like. Consider a few possibilities.

First, our decisions might be literally undetermined. They might, that is, be indeterministic events in our brains,⁵ and although various causes might affect some aspects of our choices (for instance, by making us more likely to take one option seriously), the fact that we choose one alternative rather than another might be no more susceptible of any sort of causal explanation than the fact that a uranium atom emitted one electron rather than another at a given time. In this case, we would not need to worry about the threats to our freedom posed by causal determinism. We would, however, need to worry about another problem: namely, that it is not clear why, in this case, we would be responsible for what we do.

Normally, we think that our choices are under our control. We think about what we ought to do; we weigh the considerations we are aware of, and try to think of things we might have overlooked; and then we choose accordingly. Even when we decide to do something we know we should not do, we normally suppose that we could have made a different choice had we tried harder to resist. This is why we think that we are responsible for the choices we make, and that our actions reflect something about us.

If, however, our choices are truly indeterministic, then we ourselves cannot control them; indeed, it is not clear why we should say that we choose, and not that our choices just happen to us. This makes it hard to see why we should hold ourselves responsible for what we do. When someone does something wrong, we normally blame her because we think that her conduct was under her control. But if we found out that the action was the result of an indeterministic event in her brain, then it is not clear why we would think this. She might not be the sort of person who normally does this sort of thing. She might have been trying to resist the temptation to do it on this occasion. Her efforts might be the sort of efforts that would normally allow her to resist. And yet on this occasion, they were not sufficient, not because she herself failed to try hard enough, but because her choice was indeterministic, and therefore no amount of effort on her part could possibly ensure that she did the right thing. It is hard to see why it is her fault that this time, despite her best efforts, her decision happened to turn out the wrong way.

One of the main reasons why our freedom matters to us is that we normally think that we must be free agents if we are to be morally responsible for what we do. If the idea that our choices are indeterministic undermines our moral responsibility, then it is less appealing as an account of our freedom. Indeed, it is not clear why we should regard a person whose choices were indeterministic as free, because she would not be able to control her own choices, or to try to ensure that she lived the kind of life she saw as best.

If the idea that our choices are indeterministic natural events seems unpromising, we might consider instead the idea that our choices are made by the sort of *x* factor that Gazzaniga mentions: “the ‘ghost in the machine,’ the soul, the mind, or the spirit.” Unfortunately, this option also faces difficulties. For one thing, it seems to make our freedom turn on our possession of an ability to defy natural laws, an ability that is hard to explain convincingly, and even harder to show that we have.

More importantly, however, the idea that our choices are made by an *x* factor does not solve the real problem. What threatens our freedom, on Gazzaniga’s view, is the possibility “that we live in a predetermined world—whether it’s caused by fate, preordination, or genetic hard-wiring—where every action, human and otherwise, is inevitable.” The problem Gazzaniga sees with being free in such a world is not that in it we could explain why we make the choices we do as the result of natural causes, or that we might explain them scientifically. It is the possibility of giving any causal explanation of them at all.

Because the problem Gazzaniga is concerned with does not depend on the fact that our choices can be explained in natural or scientific terms, invoking nonnat-

ural, nonscientific beings like spirits or souls will not solve that problem. If these beings operate according to causal laws—if, that is, we can explain what spirits or souls do by citing antecedent causes that produce spiritual effects according to causal laws—then the problem remains. And if the actions of these beings are indeterministic, then we face the same problems about moral responsibility that we would face if our choices were indeterministic natural events. Invoking souls or spirits simply transfers the problems presented by scientific explanations of brain activity to a different and less familiar set of entities; it does not solve them.

The fundamental problem is as follows. Gazzaniga and Libet both assume that if our choices are causally determined by something, then we are not free. If they are right, then claiming that our choices are due to something nonnatural, like a ‘ghost in the machine’ or a spirit, cannot solve the problem unless that claim allows us to argue that our choices are not causally determined by anything. If, however, our choices are not causally determined by anything, then our choices are not under our control and we are not morally responsible for what we do. Moreover, the idea that our freedom consists in a tendency to make decisions in an indeterministic, uncontrollable way makes it unclear either that we are free, or why we would want to be.

AN ALTERNATIVE ACCOUNT

Because the alternatives to determinism seem so unpromising, philosophers have tried to come up with some way of understanding human freedom that does not turn on the truth or falsity of determinism, but on our capacity for self-government—our ability to decide for ourselves what kinds of actions to perform and what kinds of lives to lead. What follows is a sketch of one such account.⁶

Suppose I knew that all my choices, and indeed my whole mental life, were caused by events in my brain. Neural causes determine what I think about, what considerations occur to me when I try to figure out what to do, which ones I take seriously and which I disregard, the strength of all my various motivations, and what, as a result, I will choose to do. How should I respond to that knowledge?

One answer might be that I should stop trying to govern my own actions and just let nature take its course. If everything is determined anyway, then I cannot affect what is going to happen, so why should I bother to try? This would be a mistake, much like thinking that either there will be coffee in my coffeepot tomorrow morning or there will not, and it is already determined which it will be, so it would be a waste of energy to bother making the coffee. After all, if it is already determined that there will be coffee, then I do not need to bother making any; and if it is already determined that there will

not be any coffee, then my trying to make some is just so much wasted effort.

Plainly, this line of thought overlooks the fact that if coffee is to be made in my house, it will be made as a result of my efforts. If it is determined that there will be coffee tomorrow morning, then that does not mean that I do not have to make any; it means that it is determined that I will make some. For that reason, if I want there to be coffee tomorrow, I should grind the coffee beans, put water in the coffee machine, and turn it on. If I do that, I am not engaging in efforts that determinism has shown to be unnecessary, but doing what needs to be done in order to have my coffee.

Likewise, the fact that it is determined that I will do something does not mean that my choices and my efforts are irrelevant to its occurring. It might be that it is determined that I will make those efforts, and that it is precisely because I make them that I end up doing what I do.

The fact that what I will end up doing is determined would imply that my choices and my efforts are irrelevant only if my choices and efforts were not themselves natural events. For if natural events fully determine what happens next, and my choices were not natural events, then obviously they could play no role in determining my conduct. But because my choices, my efforts, my deliberation, and my mental life are part of the natural world, determinism does not imply that they are irrelevant to what ultimately happens, nor that I do not need to bother with them.

In fact, determinism alone does not imply anything about what I should do next. Perhaps I will try to figure out what I ought to do, try as hard as I can to do it, and succeed; if so, then that is what it is determined that I will do. Perhaps I will fail despite my best efforts; if so, then *that* is what it is determined that I will do. Perhaps it is determined that I will conclude, wrongly, that I do not need to make any effort to do anything, and as a result I will not do what I think best. Determinism does not imply anything about which of these things is true, nor does it imply that I have any particular reason to choose one of these alternatives.

For this reason, the knowledge that all my choices are fully determined by antecedent causes would leave me exactly where I was before: with choices to make and a life to lead, and with no more reason to choose any particular option than I had before. From the point of view of a person trying to decide what to do, determinism is irrelevant.

Determinism would affect my choices if it allowed me to know, in advance, what I would do. For obvious practical reasons, neuroscience is unlikely to enable us to predict our own actions with certainty. To make such a prediction, we would need enough scientific knowledge of the workings of the brain to allow us to predict the

outcome of a given choice, a complete understanding of the state of a person's brain and body at some initial point, and knowledge of all the causal influences that would affect her between that point and the choice we are trying to predict, not to mention enough computing power to allow us to use this information to generate a prediction. This is, to put it mildly, unlikely.

It is not just impractical, however, but rather it is impossible for a person to learn what some deterministic theory implies that she will choose to do if that information would affect what she chooses. To see why, imagine that all the practical problems described above have been solved. We have developed a flawless scientific theory that allows us to predict, given information about a person's state at some earlier time and about the causes to which she is subject, what she will choose to do. Moreover, we have perfect information about some person's state at some earlier time, and all the computing power we need. All we need to do is to factor in the events that will affect that person between that earlier time and when she makes her decision, and we can predict what she will choose to do.

Now suppose that, having made such a prediction, we decide to tell that person what we have predicted that she will choose to do. In so doing we would alter her brain state in a way that our original prediction did not take into account; and so that 'prediction' would be true not of the actual world, in which we tell her what she is about to do, but of a merely hypothetical world in which we keep silent. A statement about what would have happened had things been different is not a prediction of what will happen, things being as they are.

Suppose that we try to compensate for this by making a new prediction that takes what we plan to tell her into account. To do so, we would have to know what we are going to predict that she will do in order to factor it into our calculations. That is, we would have to know the result of our calculations in order to arrive at that result in the first place.

This will not be a problem if what we tell her does not affect what she goes on to do. This might be true for several reasons. First, what she does might not depend on her at all. If, for instance, she has fallen off a cliff and we yell out to her the exact location where she will land, her trajectory will not be affected by this knowledge. Second, she might decide not to let our predictions affect what she chooses to do, and succeed. In either case, our prediction would not affect her future conduct, and therefore we could predict it. But if, in informing her what she will choose to do, we affect what she will do, then we would need to know what we will tell her in order to arrive at a prediction in the first place.

This means that a person can learn what it is determined that she will do only when that knowledge does not affect what she will do, either because what she will

do is not up to her in the first place, or because she does not allow this knowledge to affect her. If what she does would be affected by her learning what we predict that she will do, then nothing she learns could in fact be a true statement of what it is causally determined that she will do. Once again, from the point of view of someone trying to decide what to do, determinism is irrelevant.

FREE CHOICES

The fact that determinism is irrelevant to someone who is deciding what to do suggests that we might use the point of view of such an agent to construct an account of freedom. So: suppose that I am deciding whether or not to take a walk, that determinism is true, and thus that causes now exist that fully determine whether I will take a walk or not. Suppose further that if I choose to take a walk, then I will take one, and if I choose not to, then I will do something else, and that I am not agoraphobic, compulsive about walking, or subject to any other psychological condition that prevents me from choosing whichever course of action I want. I can choose whatever I please, and I will do whatever I choose. The only threat to my freedom comes from the fact that my choice itself is causally determined.

In this situation, I am deciding between two alternatives: two courses of action that I can perform if I choose to do so. I take the question of which alternative I will choose to be open, not because I believe that determinism is false, or that my choice is uncaused, but because I take that question to be one whose answer depends on me, and which I have yet to answer. Because I take that question to be open, I should not, for these purposes, think that my two alternatives differ with respect to their possibility. I can perform either alternative if I choose to do so, and I have not yet decided which I will choose. All that prevents me from performing either of them is my own decision to do something else.

If determinism is true, then that decision is itself the result of antecedent causes—but that fact is irrelevant to me when I am deciding what to do. For one thing, it is impossible in principle that I know what in particular it is determined that I will choose, if that knowledge would affect my deliberation. Therefore I cannot avoid the need to decide for myself what to do by simply embracing the inevitable. For another, when I try to figure out what to do, I am not trying to figure out what causes will ultimately lead me to choose one way or another; rather, I am trying to figure out what reasons I have for choosing one alternative over another, and the fact that my decision is itself determined is not relevant to this question, any more than it would be relevant if I were, say, trying to figure out the answer to a mathematical problem. It is both unavoidable and rational that I view the question of what I should do as one that it is up to me to answer,

whether or not the answer I give is itself causally determined.

When our actions are up to us in this sense, I would argue that we are free. This does not mean that we have to act as if our choices are uncaused, or that we have to accept the illusion of genuine freedom for practical purposes. For one thing, I have not argued that, when we try to figure out what to do, we have to believe that our choices are not determined; indeed, I see no reason to believe that this is true. I have argued only that the truth or falsity of determinism is irrelevant to us when we make decisions.

More importantly, however, the point of these arguments is not to show that we cannot help thinking that our choices are uncaused even when they are not. Rather, it is to replace the idea that freedom requires that our choices not be caused with a different conception of freedom, one that holds that we are free when we can do what we choose, and thus when we rightly regard our conduct as up to us. Whether or not we are free in this sense is not called into question by the truth of determinism, nor by discovery of the neural mechanisms that govern our minds. If this account of freedom is correct, then we are in fact free when what we do is up to us.

CONCLUSION

On the account I have sketched, freedom requires not causal indeterminism, but a capacity for self-governance. When we can decide for ourselves what kind of life to lead, what kind of person to be, and what kinds of actions to perform, then we are free. Our freedom can be under-

mined in many ways: by phobias and compulsions; by failures of self-control, of empathy, or of the capacity for planning and execution; by mania or psychosis. Neuroscience, by helping us to understand these conditions and the neurological problems that underlie them, holds out the promise of illuminating the mechanisms by which such factors undermine our capacity for self-government.

In so doing, advances in neuroscience might allow us to discover ways to help those who suffer from such conditions to regain full control over their lives. They might also lead us to conclude that some persons whom we had always regarded as free agents are not, by disclosing some unsuspected pathology that prevents those persons from effectively exercising a capacity for self-governance. Advances in neuroscience should not, however, lead us to doubt that any human being has ever been free.

As long as we have the capacity to decide what kind of life to lead and what kinds of actions to perform, and to act on our decisions, we have freedom of the will, whether or not our decisions have neural causes.

REFERENCES

1. Gazzaniga M. *The ethical brain*. Washington and New York: Dana Press, 2005.
2. Libet B. Do we have free will? *J Conscious Stud* 1999;6:47–57.
3. Hobbes, Thomas: *Leviathan*. Curley E, editor. Indianapolis, IN: Hackett, 1994.
4. Hume, David: *An enquiry concerning human understanding*. Beauchamp TL, editor. New York: Oxford University Press, 2006.
5. Kane R. *The significance of free will*. Oxford: Oxford University Press, 1996.
6. Bok H. *Freedom and responsibility*. Princeton, NJ: Princeton University Press, 1998.