



# Defining free will away

EDDY NAHMIAS ISN'T ASKING FOR THE IMPOSSIBLE

*Free Will* by Sam Harris  
(The Free Press), £6.99/\$9.99

In his pithy pamphlet *Free Will*, Sam Harris explains why he thinks free will is an illusion and why this matters. Given his other books, one would expect science to drive Harris's conclusions, but here his argument is conceptual. Step 1: Define free will in such a way that it is impossible. Step 2: Remind us that we cannot have what is impossible. My response is just as simple: Harris's definition of free will is mistaken. To have free will, people don't need the impossible; nor do most people think free will requires the impossible.

Harris never provides a clear statement of his argument, but it relies on defining free will as requiring some nebulous X-factor: (1) Free will requires X; (2) X is impossible; So (3) free will

is an illusion. Harris gestures towards several such X-factors: having "an extra part of me" that transcends brain and soul; being free to "do that which does not occur to me to do"; being unpredictable in principle; not being " beholden to the laws of nature"; and the more reasonable condition that "we are the conscious source of most of our thoughts and actions". He also says free will requires that we "could have behaved differently than we did in the past", but he does not focus on this condition and ignores the vast philosophical literature on it. Harris concludes: "Consider what it would take to actually have free will. You would need to be aware of all the factors that determine your thoughts and actions, and you would need to have complete control over those factors."

## Harris's definition of free will is mistaken

To the extent that one can make sense of these X-factors, most of them clearly *are* impossible – people can't have magical powers to transcend themselves or the laws of nature, nor can they create conscious thoughts from scratch or know everything about why they think or act as they do. So, all the work occurs in premise 1. How does Harris defend the definition of free will required get the argument off the ground? Like other scientists who proclaim that free will is an illusion, he simply asserts that it captures “the popular conception of free will”, without offering any evidence to back up this claim.

Harris misreads or ignores the arguments offered by the majority position in philosophy, compatibilism, which rejects the impossible conditions Harris foists on free will. His response: “the ‘free will’ that compatibilists defend is not the free will that most people feel they have.” Harris mistakenly describes Daniel Dennett's view as suggesting that conscious agency is irrelevant to free will and then uses this misrepresentation to represent *all* compatibilist views, including ones that explicitly discuss Harris's concern about how we can consciously control our actions in light of competing desires. He quotes my *New York Times* column describing free will in terms of having capacities for conscious deliberation, planning and self-control and having the opportunity to exercise these capacities in action. Harris responds that “these phenomena have nothing to do with free will”. Nothing?

It's ironic for a scientist to assume from the armchair something that could be studied

empirically – what people actually think about free will. I don't have space to substantiate my response here, but unlike Harris, I am *not* making it from the armchair. Along with others, I have done experimental work exploring ordinary intuitions about free will and responsibility. For instance, we have studied whether people think free will would be illusory if neuroscientists could use a brain scanner to predict everything people would do before they are aware of deciding what to do. Harris uses this idea to illustrate what he assumes everyone thinks about free will, but his assumption is mistaken. Our data show that, as long as the neuroscientists don't manipulate people, the vast majority of people do *not* take such predictability based on prior brain activity to threaten free will (see box p. 112). Our other results suggest that most people have compatibilist intuitions about free will and determinism, as long as they don't make the sort of mistake Harris both tries to avoid and falls into – the mistake of thinking that our conscious mental activity is causally bypassed if it is part of the natural order.

Indeed, if I were to analyse Harris's underlying worries about free will, I'd say that his main concern is that if our brains do it, *we* don't. Harris suggests that the conscious capacities for deliberation and decision-making associated with free will are simply along for the ride, observing what our brains have already done: “I, as the conscious witness of my experience, am not the deep cause of it.” I'm not sure what he means by “the deep cause”, but I suspect that when you consciously plan how to get twelve

## Experimental Philosophy on Free Will

Sam Harris writes, "If we were to detect [people's] conscious choices on a brain scanner seconds before they were aware of them ... this would directly challenge their status as conscious agents in control of their inner lives". Here is a scenario we have run that is very similar to Harris's thought experiment. Participants read the following case (slightly abridged here):

Recent brain scanning studies have shown that specific types of brain activity can predict what decision a person will make a few seconds later, before the person is even aware of having made a decision. Now, imagine that in the future this technology becomes much more advanced so that a neuroscientist can use a brain scanner from a distance to obtain information about a person's brain activity without that person knowing it. The neuroscientist can use that information to predict what the person will do and can even use the scanner to alter the person's brain activity so that they will do something else. Imagine that Mr Jones is in a voting booth where he is considering whether to vote for either the Republican or the Democratic candidate for President. Dr Black is a neuroscientist who, without Jones knowing it, is using his brain scanner to monitor Jones's brain states and alter them if need be. Dr Black has enough information to know that if the brain scanner indicates that Jones is in brain state D22, then he will vote for the Democrat, and if the scanner indicates that Jones is in brain state R16, then he will vote for the Republican. Dr Black is a diehard Democrat, so if the scanner detects any brain activity other than D22 indicating that Jones will *not* vote for the Democrat, then Black will activate the scanner to alter Jones' brain activity before Jones is even aware of making a decision in order to ensure that Jones votes for the Democrat. As a matter of fact, Jones's brain activity indicates that he will vote for the Democrat, so Dr Black does not do anything. Jones votes for the Democrat without any intervention from Black or his brain scanner.

Participants then rated their agreement or disagreement with statements, including these:

- Jones voted of his own free will.
- Jones is responsible for how he voted.

Nearly 70% agreed with both statements (with about 20% disagreeing and 10% neutral).

This despite the fact that a minority agreed with: "Jones could have chosen *not* to vote for the Democrat". The majority also responded yes to this question: "Do you think that it is possible that, in the future, there will be a technology that will allow scientists to predict what people do, as suggested in the scenario?"

## Harris never provides a clear statement of his argument

complicated tasks done before lunchtime and you exercise self-control to avoid quitting some of them, your conscious mental activity is playing a causal role in how your body moves. Do your conscious thoughts have prior causes, perhaps fully determined and surely beyond your complete understanding? Of course. But causes can be caused. And if your conscious thoughts are an integral *part* of the causal stream *because* they are processes occurring in your brain, then you are not just observing what your brain has already done – your conscious thinking and efforts of self-control have downstream effects.

I suspect that Harris and most other scientists who are sceptics about free will fall into this trap of thinking consciousness is a passive bystander because they find themselves in limbo about how to understand consciousness. Steeped in the modern zeitgeist, they believe that neuroscience will surely explain all human decisions and actions. But such explanations currently offer no place for consciousness, since there is not yet a neuroscientific theory that explains how certain neural processes are the basis of and explain conscious processes. So, they fall into the trap of thinking that your conscious self, if ultimately based in the brain, plays no role in your decisions and actions: as Harris says, “your brain has already determined what you will do”. If you can’t really get yourself out of the very dualist picture you are trying to reject, you end up saying things like Harris does: “The choice was made for me by events in my brain that I, as the

conscious witness of my thoughts and actions, could not inspect or influence.”

The solution here is to recognise our place in history. Neuroscience is a young science. As it matures, consciousness and rationality will be the hardest mental phenomena for it to explain. However future neuroscientific theories explain conscious and rational processes, it is highly unlikely they will show that the neural processes involved in conscious deliberation, rational planning, and self-control are somehow side-streams that have *no* downstream effects on what we do. If these processes *are* a part of the stream, this will explain how we are “authors” of our choices, since authors can create even if caused to do so. Harris *nearly* recognises some of these points, but then gets twisted into knots worrying about our thoughts coming “out of the darkness of prior causes”, the conscious part of the mind “living at the mercy of other parts”, and being unable to “*choose* to choose what I choose”.

It is ironic for scientists like Harris to think that science will explain *away* free will, rather than helping to *explain* how it works. Organic chemistry did not make life disappear by explaining how living processes work. Copernicus did not explain *away* the earth by explaining how it moves. The only reason to think free will can’t be explained is to define it such that it *must* be inexplicable – to assume that people demand the impossible.

Harris is not worried about any negative effects his book will have on the (many) people

## He misreads or ignores the arguments offered by the majority

who will read it: “Speaking from personal experience, I think that losing the sense of free will has only improved my ethics.” But here again, a sample size of more than one would be useful. Scientific research can help to determine what people believe about free will, what they think would threaten it, and how they feel and act when these beliefs are threatened – for instance, when scientists proclaim free will is an illusion. Such research is now occurring. The next step will be to study how such information might influence our legal and political systems, in part so we can use such information to improve these systems.

Here, I agree with Harris that what people think about free will matters. I also agree with him that our society should tone down its over-inflated sense of “entitlement to the fruits of [one’s] good luck” and its harsh retributive impulses towards criminals and those down on their luck. But these goals, I predict, are not best met by proclaiming that *no one* has *any* free will and that no one *really* deserves credit or blame for *anything*. Instead of defining free will as all-or-nothing and concluding that no one has any, better to recognise that people have *degrees* of free will, depending on their capacities for conscious and rational self-control and on their opportunities to exercise these capacities. This view accords with the way we actually think about children coming to have more and more responsibility for their actions as they mature. It accords with scientific discoveries regarding

our limited resources for self-control and with discoveries regarding the impact specific genes, environments, or brain disorders can have on one’s capacities for rational self-control. Harris is wrong, however, to suggest that “the more we understand the human mind in causal terms, the harder it becomes to draw a distinction between” behaviour caused by a brain *tumour* and behaviour caused by one’s *brain*, as if there is no distinction between abnormal and normal functioning of the brain and hence the mind.

When Harris proclaims that free will is an illusion, people will likely interpret that to mean they lack what *they* think free will is. People think free will includes conscious deliberation and self-control. But we *don’t* lack these capacities, and it won’t improve our lives or society if we start thinking we do. Unlike the impossible self-creation and self-knowledge Harris foists upon free will, a more reasonable and accurate understanding of free will is amenable to scientific study. Science is likely to show we have *less* free will than we tend to think, and learning this may move us towards Harris’s practical goals. But a scientific understanding of the mind need not mislead us into believing we lack conscious agency or that we are just, in Harris’s words, “calibrated clockwork”.

**Eddy Nahmias** is associate professor of philosophy and neuroscience at Georgia State University. He is writing *Rediscovering Free Will: Autonomy and Responsibility in the Age of the Mind Sciences*.