

Greene and Cohen: Neuroscience and free will

THE NEW NEUROSCIENCE. Here we limit our concern to the subjects related to metaphysics, rather than to ethics or law. That said, some of these subjects ought to be mentioned.

Neuroscience, according to Greene and Cohen, has no effect on the law in the sense of not aiming, or being able, to alter any specific legal statutes—once the ‘legitimacy’ of the legal framework is taken for granted. But it may well put into doubt that very legitimacy to begin with. 1778.2

What this difference amounts to gets clearer when Greene and Cohen argue that people’s intuition about moral responsibility turns on the question, ‘Was it *him* when he did it?’ The traditional legal requirement, however, is rationality: ‘Was he rational when he did it?’ And on many occasions, the answers co-incide. But not always.

In particular, it may be the case that you act rationally, but your rationality was manipulated by some external agency. In this case, Greene and Cohen argue, our intuition ‘*X* is not himself’ trumps our intuition ‘*X* is rational’.

Example 1 (Usual case). Suppose *X* shoots someone when enraged. Then I might say:

(9-1) He was not *himself*. He was so enraged he forgot himself, couldn’t think straight. (Hence he isn’t fully or at all responsible for shooting etc.)

I implicitly identify ‘being yourself’ with ‘being in control of your actions’, ‘making a rational choice’. As one condition fails, so does the other.

Example 2 (Plato’s Guardians). Plato tells us in the *Republic* that the Guardians are maximally rational, hence just. But their rationality is determined (i) by relentless horse-like breeding, (ii) by careful education from the get go that penetrates every aspect of their lives. So it is not pointless to ask: are the Guardians brainwashed? did they choose to be who they actually are? are they, therefore, free in the sense of required by full legal (and moral?) responsibility?

Example 3 (Boys from Brazil). Greene and Cohen imagine, with Gideon Rosen’s and Gregory Peck’s help, essentially a *Republic*-like scenario where people are bred and educated in the way that Hitler was. Well, when one such individual (Mr Puppet, or ‘Rudolf’) begins behaving like the actual Adolf, are we to say that he was acting freely and was responsible for what he was doing? Although his regular rationality is not compromised, our intuitions shift. Or so Greene and Cohen argue. We increasingly see Rudolf as a puppet manipulated by the breeder. In effect he is free no longer. 1779.2

The implications are stark. If Rudolf is a puppet, isn’t Adolf a puppet too? Even if there was no conscious external agency putting Adolf in motion, there were no less powerful impersonal forces like genes, the father, Austrian politics and education. If the Guardians are, we say, manipulated or brainwashed, is the most rational agent not in the same condition?

Our intuitions shift when we realise that there is some ultimate external source of your character and your motives, including your rational motives. And this means that we tend to distinguish between what *you* are and what your brain (along with other parts of the body) is or does. Now any empirical/psychological reflection would show that there is no tenable distinction there. You *are* your brain. If there is anyone in the driving seat, when you act, it must be your brain. There is no homunculus manipulating the brain and the rest of the body. 1779.1

Remark 4. Neither should there necessarily be a unified control centre in the brain that we could identify with the ‘self’ now interpreted thoroughly reductively. There may be rather different modules vying for attention, so to speak. When you are voicing an opinion, expressing an intuition or a feeling, you are acting like a spokesman of these modules (that’s Minsky’s and Kurzban’s favourite metaphor).

Remark 5 (Margalit’s gambit). Some moral philosophers argue that your human agency, and with it responsibility, may be recovered by emphasising the ability to re-evaluate, reset your behaviour from scratch. If, for example, you are an evil Unnamed from *The Betrothed*, you can at any time reverse your behaviour and become a saintly type. But this response simply begs the question. We are not convinced that *every* one is able to reset, that there are no psychological dispositions antecedently present (and causally determined) that enable the reset. Moreover, as with Unnamed, there has to be an unusual external stimulus of a very specific nature to enable the reset. Presumably the role of the stimulus is to interact with the already given psychological characteristics of the person. Absent the stimulus, your free will is simply too weak to effect a dramatic change. (*The Betrothed* has many other interesting characters to illustrate the problem of free will and moral choice.)

FOLK PHYSICS AND FOLK PSYCHOLOGY. Greene and Cohen propose to diagnose the attribution of free will to others (also the conflicts it brings on its heels). This attribution comes from two distinct cognitive modules. The folk-physical module parses the world into bodies that are subject to causality, but do not themselves carry causal powers. The folk-psychological module identifies agents, i.e. objects that are uncaused causes, prime movers, that change things. Of course, things can also change others things, but their causality always originates outside of themselves. Crucially for the ethical implications, these agents are endowed with their own internal causality—desires, volitions, emotions etc.—that may provide an explanation of analogous to the folk-physical causal explanation. Thus folk psychology is the ‘gateway to moral evaluation’.

1781.2

1782.2

The diagnosis is strikingly, if unsurprisingly, Kantian. Once we see the world with our folk-physical module, we see causality everywhere and no agency, no freedom. Once we see the world with the folk-psychological module, we see freedom at least somewhere, with some objects. To make sense of actions we need agents. So we read freedom into the minds (or the ‘minds’) of some objects, and then we are able to converse in moral terms. Note too that agency needn’t be restricted to humans or even animals: you may smash your car in anger if it fails to start, though you normally would deny it any agency. That is: most often you see the car as inanimate, though there will be fleeting occasions when you see it as active. Similarly with computerised circles and triangles. With humans, it’s the other way round.

1782.1

KANTIAN BLUES. If we contemplate the two-stance approach to agency and action that is based on two different cognitive modules, then both might seem equally legitimate. Both *are* legitimate: it all depends on our concomitant interests. So with Kant, we see others as empirical egos, so long as we reason about them as natural objects. We see them as pure egos, so long as we try to understand their actions.

But, according to Greene and Cohen, we shouldn’t be lulled into false symmetry. For we must conclude that free will is an illusion produced by our cognitive architecture. Well, why?

1784.2

Here recall, first, that Kant (as I understand him) didn’t mean to say that the two stances are equally legitimate descriptively. In our descriptions of the world we should be uncompromising causalists (i.e. ‘determinists’, but as Greene and Cohen helpfully note, we must allow here quantum indeterminacies, so the label is misleading). It is only when we reason about moral facts—and such facts are just given to us!—that we must adopt the libertarian stance. The two stances are not on the same footing even according to Kant.

1780.2

In any event, Greene and Cohen’s overarching framework is not Kantian to begin with. They begin explicitly from within empirical science—that is, within causalism. The operations of the folk-physics module are congruent with this framework. But the operations of the folk-psychology module are a weird outlier. That is the main reason, I believe, why we must judge them as second-rate, and the attribution of free will as an ‘illusion’.

YSB