Hume's metaphysics (2): Free will (*EHU* section 8)

The view that each of us is a free agent who, under normal circumstances, makes genuine choices for which she can rightly be held responsible, and thus for which she can be praised or blamed, seems to be written into the very idea that we have of ourselves as persons. Moreover, our practice of holding one another responsible for what we do is the foundation of our organised systems of reward and punishment, and thus of our legal codes. So if we were to discover that we don't have free will, we would, to put it mildly, have lost a keystone of the world as we see it.

Some philosophers (e.g., John Searle, 'The Freedom of the Will', in his *Minds, Brains, and Science*) have argued that what sits at the core of our conception of free will is an idea that we can call the *could-have-done-otherwise* principle. Consider some action performed by a person. If it is true to say of that person that she *could have done otherwise*, all other conditions remaining the same, then, on that occasion, she exercised free will. Prima facie, there seems to be something right about this idea, since the *feeling* that we could have done otherwise than we in fact did is surely a ubiquitous feature of our *experience* of making choices and of performing actions.

Hume on free will

In *EHU* 8*,* Hume considers the free will issue. The key to understanding Hume's discussion here is the fact that he is a *determinist*. Determinism is the philosophical position according to which all events, *including human actions*, are determined by prior causes. For the determinist, the occurrence of every event, including every human action, is causally determined by the prior occurrence of some antecedent event. Hume suggests that there is nothing in determinism that conflicts with common-sense. Indeed, he argues that since we all believe our actions to be the result of inner causes (e.g., our motives and character traits), inner causes which are often identifiable, determinism represents the common-sense view.

Hume's determinism

Let's look at Hume's determinism in more detail. For Hume, the connection between human motives and actions is essentially the same as that between non-psychological causally related natural events. So let's recall the key features of Hume's theory of causation. In analysing section 7 we distinguished between a metaphysical reading of Hume on causation, and an epistemological reading. On the metaphysical reading, Hume holds that there is no necessary connection between causally related events, there is only regular succession (constant conjunction plus the invariable temporal ordering of the cause-event before the effect-event). It is this interpretation that fits best with Hume's

occasional talk of necessity being something that exists in the mind, not in objects. On the epistemological reading, Hume holds that while there may well be necessity in nature, *all that we have knowledge of* in causal sequences is regular succession. On either reading (metaphysical or epistemological), we form the idea of a necessary connection between causally related events due to the operation of the psychological mechanism of custom, plus the feeling of inevitability that subsequently accompanies our inductive inferences from causes to effects. In *EHU* 8, Hume applies this account of causation to the links between human motives and human actions. He holds that there are observable constant conjunctions between motives (including character traits) and actions in human behaviour, and that particular kinds of motives have been observed always to come before particular kinds of actions. Thus, he claims that if one looks at human history, one discovers that the 'same motives always produce the same actions' (*EHU* 8.7).

The Humean world, then, is a *law-governed* place: given the deterministic character of the world, it would violate a law of nature for a prior event to occur and a subsequent event not to. (On the metaphysical reading of Hume, a law of nature turns out to be nothing over and above a universal regularity, a true universal generalisation.) For Hume, this picture of the universe as law-governed applies just as much to human actions as it does to the behaviour of the rest of the natural world, so we have the grounds for a genuinely scientific psychology of an appropriately Newtonian kind.

Problem about the indeterminacy of psychological laws

About now a potential problem with Hume's view becomes visible, namely that any supposed universal laws about human behaviour that we care to state seem always to be false, because they have exceptions. So, for example, let's say that you have been observing my bar-related behaviour, and you have concluded that my movement towards the bar is invariably the result of my desire to obtain a beer, and my belief that I can get one at the bar. You might think that this is a universal piece of human psychology. If so, you would be moved to say that for all X (where X is any person), if X desires a beer, and if X believes that s/he can get one at the bar, then X will go to the bar. This law of human psychology will soon run up against exceptions. Maybe X is too mean to buy a beer, or perhaps s/he has enjoyed such a large number of visits to the bar already that s/he can no longer move. You get the idea. Lots of factors can intervene that would produce exceptions to, and thus falsify, the proposed universal law. A natural first reaction here is to include a list of exceptions in one's statement of the law. (So …X will go to the bar *unless* X is too mean, X is too drunk, and so on…) But any such list looks to be open-ended and impossible to complete. So perhaps the solution is to add in a *ceteris paribus* (all things being equal) clause, so that the law reads: for all X, if X desires a beer, and if X believes that s/he can get one at the bar, then, *all things being equal*, X will go to the bar.

Some philosophers think that this move stinks, since it boils down to saying that X will behave in such and such a way unless X doesn't, which is to say precisely nothing. Laws that say nothing can hardly form the basis of a genuinely predictive or explanatory scientific psychology.

Perhaps the proper Humean response to this complaint would be to point out that things aren't actually all that different when it comes to non-psychological natural events, and we don't conclude that the relevant laws of nature there are vacuous. Thus Hume observes that some natural events appear to be exceptions to our scientific laws, but he claims that 'almost in every part of nature, there is contained a vast variety of springs and principles, which are hid, by reason of their minuteness or remoteness', and that it is the 'secret operation' of these underlying causal processes, rather than any 'contingency in the cause', that accounts for the apparently exceptional phenomena (*EHU* 8.13). In other words, although it may sometimes look as if determinism doesn't hold in the natural world, that's only because our science hasn't yet penetrated the workings of nature far enough to identify the truly exceptionless, low-level causal laws that really govern events. (Think of something like a completed atomic physics.) If we had such a science, events that constitute exceptions to our higher level laws would nevertheless be explicable using, and so would not contradict, our truly exceptionless lower level laws. The laws of the special sciences (such as geology) are thus high-level approximations to the truly exceptionless laws that describe the low-level causal processes actually in play.

On this view, the general psychological laws that we discover by observing constant conjunctions between human motives and behaviour are a species of such high-level approximations. This idea can be developed in such a way that the conceptual distance between Hume and many contemporary determinists is minimised. One place in which deterministic causal links might conceivably be found is at the level of neurophysiologically identified brain-states and physical bodily movements. Indeed, given contemporary science, it does seem plausible that if I were in exactly the same brain state as I was 10 minutes ago, in exactly the same surrounding conditions as I was 10 minutes ago, then I would engage in exactly the same bodily movements as I did 10 minutes ago. Hume doesn't formulate his determinism in these terms, but it would, in many ways, be consistent with his position to do so. And contemporary determinists often do state their position at the level of brain-states and bodily movements. So once determinism is posed in these terms, it seems that the view on offer is not only Hume's; it may well be ours too! Our contemporary scientific world view is also one in which the universe is a giant causal system, unfolding deterministically according to the exceptionless causal laws that would be found in a completed physics.

The free will problem

This leads on to what is in this context the main concern. For the determinist, every event is causally determined by some antecedent event. So if the universe were to be rewound, so that exactly the same set of conditions as previously produced some event are once again in place, then exactly the same event would happen again: same cause, same effect. If each of us is part of that natural order, then there seems to be no room for the universe to unfold in a different direction to the one which it has, in fact, taken. In other words, determinism seems to be inconsistent with the could-have-done-otherwise principle, the principle which, as we saw, appears to be at the conceptual heart of our notion of free will. So determinism looks to be *incompatible* with the freedom of the will. And if this apparent incompatibility is genuine, then determinism is also in conflict with our established social practices of praise/blame and reward/punishment.

The problem here can be stated in the following way. Logically speaking, there are four options:

1. Determinism is false, and there is free will.

2. Determinism is true, and free will is an illusion.

3. Determinism is false, and free will is an illusion.

4. Determinism is true, and there is free will.

1, 2, and 3 are unpalatable: 1, because the fact that the universe is deterministic is part of our accepted scientific world-view; 2, because we want to preserve our vision of ourselves as free agents; and 3, because of both of those reasons. We'd like to endorse 4, but so far it seems that that option is just not viable. Hume (like others before him, and like many since) however does in fact pursue option 4, an option which has come to be known as *compatibilism*. (Hume doesn't use the term compatibilism. He says that he is engaged in a *reconciling* project (EHU 8:23).)

Hume's compatibilism

At this point there occurs a (by now) familiar Humean move, namely a demand that we define the key terms in the debate clearly, coupled with a prediction that this will enable us to see how the problem in question may be resolved (*EHU* 8:23). In the present debate the key terms, according to Hume, are *necessity* and *liberty*. On Hume's view, the correct understanding of *necessity* is, of course, provided by his theory of causation (see above). But the pivotal manoeuvre in *EHU* 8 is where

Hume defines *liberty* as 'a power of acting or not acting, according to the determinations of the will' (*EHU* 8:23). So, on Hume's view, an action counts as free (and thus qualifies the agent for praise, blame etc.) when (a) the person wills it, and (b) when there are no external, constraining conditions which force that person to do other than what she wills. Put crudely, and in a somewhat extreme form, the idea is this: if I'm in a locked cell, tied up with chains, then I'm not free, because even if I will myself to do something useful such as run away, I'm simply not able to execute my decisions.

The important point to notice about Hume's definition of liberty is that it in effect denies that having free will turns on our being able to answer 'yes' to the could-have-done-otherwise question (see above). Nothing about the two defining conditions introduced by Hume is inconsistent with the following thought: if we recreate exactly the same set of circumstances now, as previously produced some *freely willed* (by Hume's definition) action by an agent – and that includes the same motives, character traits, beliefs, desires etc. on the part of that agent – the freely willed action performed by that agent now, would be exactly the same as the freely willed action performed by that agent previously. In other words, as a definition of free will, it seems to allow for an action to be both freely willed *and* causally determined. Indeed, Hume goes further. He argues that if one accepts the principle that we have been calling could-have-done-otherwise principle, then one is committed to saying that an act of will counts as free when it is *uncaused*. And that position, according to Hume, undermines, rather than secures, the idea of moral responsibility, since for an agent to be truly responsible for an action, that action must flow causally from her 'character, passions and affections' (*EHU* 8:31). It cannot be uncaused.

Hume's analysis of what is required for responsibility also helps us to see why the way to make conceptual space for the freedom of the will is most certainly *not* to deny determinism by appealing to the existence of chance events in the world. In contemporary debates about free will, some thinkers have made this kind of move by appealing to the idea that determinism is false at the quantum level, where (on some accounts of quantum physics anyway) genuinely random events may occur. But even if such chance events do occur, and even if they do have an effect on the world beyond the quantum level, the fact remains that they are surely events over which our wills *have no control*; so it is hard to see how they could establish the existence of events for which we are responsible, and thereby secure the freedom of the will.

Some modern philosophers who endorse determinism (e.g., John Searle) think that Hume's definition of liberty is inadequate. They remain convinced that any decent account of free will must appeal to something like the could-have-done-otherwise principle. The problem remains live.