

Free Will versus Natural Necessity?

Stuart Greenstreet asks if they can be reconciled

Immanuel Kant put a challenge to later philosophers which so far none has been able to meet. That we are free to choose is, he thought, a truth so intuitively self-evident that no one could argue our freedom away. Nor could we ever believe there's a true contradiction between our freedom of action and the necessary application of the laws of nature, for we can't give up the idea of nature any more than that of freedom. So even if we're never able to work out how our freedom is possible, at least this apparent contradiction ought to be convincingly removed. For if it is ever shown that belief in our freedom contradicts nature we would have to stop believing we're free and accept that all our actions are at root causally determined.

We believe that we are free to choose and that we are subject to nature's causal laws. Hence Kant's challenge to us: please reconcile these apparently contradictory beliefs. He himself insisted that we can only conceive a rational person as free if we regard his actions as not being determined by causes "external or alien to himself." So he came to believe that natural necessity as it applies to human actions must be merely *instrumental* – a tool of the rational will. We do not possess a power of free choice as beings that are subject to causal determination, he concluded – we possess it as we 'supersensibly' are in ourselves. Our power to exert "reason's peculiar causality" (i.e., free will) is a proof of "the supersensible in the subject," a realm that is in every other respect beyond our ken.

Kant-style free will is inescapably supernatural (or certainly non-natural) property. But how, we are compelled to ask, could man's causal power – a natural phenomenon – be activated by this mysterious supernatural lever? Kant replied that it is pointless to ask, for the causality of reason "is thought under freedom, in a manner that is not further or otherwise explicable." A response such as this fails to satisfy the secular mind, which expects any answer to stay within the bounds of nature, and therefore of science.

Kant's response is *dualist*, meaning that since human beings are both rational and animal they are composed of *two substances*. No response to Kant's challenge can avoid employing this elusive concept of 'substance'. A substance is a thing that exists independently of other things and is the bearer of properties. Hence we may say of a certain substance that it 'is extended' (occupies space) or 'is sensible' (detectable by the human senses). A thing's substance is what it *essentially is*, and it is not itself a property of anything else. ('Essence' is indeed a synonym for 'substance'.) In his *Meditations on First Philosophy* (1641) René Descartes proposed the dualist theory that substances are either material or mental, neither type being reducible to the other, so that no mental property (e.g., 'is a thinking thing') could ever be explained by any of our physical properties. So the reasoning that goes on in our minds could

not be explained in terms of neural activity in the brain. Kant's response to the problem of free will is also of this dualist kind, and therefore has to rely on the supernatural in order to explain the freedom of the rational will.

An explanation of free will that did not appeal to the supernatural (to anything outside of nature) would have to be 'ontologically monist', meaning that it would have to assume that human beings consist of just *one* substance. Such a monist theory was proposed by the American philosopher Donald Davidson in 1970, in an ingenious essay called *Mental Events*. Davidson made clear that this was a direct response to Kant's challenge by quoting it in full at the start of his essay.

Donald Davidson: Anomalous Monism

Davidson says it's undeniable that mental events play a causal role in the physical world, and that causality requires laws. However, as he admits, there are no laws on the basis of which mental events can be predicted or explained. These become his three basic premises, which he calls the principles of (i) causal interaction, (ii) the nomological (or law-like) character of causality, and (iii) 'the anomalism (*sic*) of the mental'.

Davidson's example of the causal interaction between mental and physical events is the sinking of the *Bismarck*: "If someone sank the *Bismarck*, then various mental events such as perceptions, notings, calculations, judgements, decisions, intentional actions and changes of belief played a causal role in the sinking of the *Bismarck*." However, as we saw, his three principles contain Kant's 'apparent contradiction': the first two imply that mental events can be predicted and explained, and the third denies this implication. Even so, Davidson adheres to all three principles. He therefore has to reconcile the fact that mental events are part of law-governed nature with the fact that they do not have law-like relations with physical events. By showing how the third principle is consistent with the other two he would also show how freedom in human action is possible. He argues that the thinking involved in mental events (such as when we 'make up our minds') is law-governed *as something else* – that is, as a physical event or process. Hence his conclusion is *monism*: mental events are tokens of physical events (and *vice versa*) and subject to nature's laws; but even so, the mental realm is both anomalous and autonomous.

Davidson's argument goes like this: We distinguish mental events from physical events by the way we pick out their respective properties. An event is mental if and only if its properties have a mental description – i.e., one that uses a mental verb to express a propositional attitude such as believing, intending, desiring, hoping, knowing, perceiving, noticing, remembering, and so on. Physical events, in contrast, are those picked out by a physical vocabulary, one which describes their attributes in

“If someone sank the Bismarck, then various mental events such as perceivings, notings, calculations, judgements, decisions, intentional actions and changes of belief played a causal role in the sinking of the Bismarck.”

terms of being objective, law-governed, and possessed of an intrinsic scientifically-discernable nature. Because all events are physical, each mental event can also be picked out using the physical vocabulary; but since there are no psycho-physical laws, no purely physical predicate (‘Her brain is in state *x*’) will, *as a matter of law*, pick out the same things as a mental predicate (‘Therefore she believes or intends *y*’). Davidson’s monism is anomalous because his thesis denies (a) that mental phenomena can be described in purely physical terms; and (b) that there can be strict laws correlating the mental and the physical. Hence the mental depends on but cannot be reduced to the physical.

Monism means the reduction of a set of things to one ontological substance or property [‘ontological’ means a fundamental category of being]. In Davidson’s theory ‘the mental’ is not an ontological category at all, but merely a conceptual one. He assumes that even though mental events are reducible to physical events (i.e., to one and the same physical process) they still retain their distinct identities as *concepts*. So he concludes that mental events are ontologically but not conceptually identical with physical events – hence monism. And it is *anomalous* monism because there are no laws on the basis of which anyone could predict or explain how events *conceived* as mental will interact with events *conceived* as physical.

Now a concept is what a person has when he or she is able to grasp some portion of their language. We understand what words or perceptions or ideas mean by subsuming them under a hierarchy of concepts. But concepts *per se* have no causal clout: they cannot cause anything themselves. If ‘the mental’ is not an ontologically sovereign thing, and mental events are merely a type of concept, then such ‘events’ could not by themselves play a causal role in the world. A consequential mental event occurs when one chooses a certain course of action, although that choice would never get beyond being a mere intention unless its physical token co-operated by executing the choice. But how could what we had in mind (our mental intention) be carried out unless mental events were somehow correlated by law with physical events? Given monism, all events are physical; so mental events = physical events. That they are not correlated by law would seem as mysterious as the existence of a causal power which, if we were to believe Kant, could produce effects in the world of the senses even though it is grounded in the supersensible.

If what Davidson gives us is simply two alternative ways of describing the same physically determined events, isn’t the mode of description which does not refer to the determination of those events (i.e., the mental description) a subservient one and, in the final analysis, redundant? For if all causal relations are subject to strict laws, and if there are no psycho-physical laws, then any instance of mind-body causation must be governed by *physical* laws. Conversely, if mental events really can cause effects all by themselves, is there not a case for saying that there are two things in question, not one?’

Kant’s dualist reply to his own challenge failed to convince us because we cannot imagine how ‘reason’s peculiar causality’ could make anything happen if one part of it (causality) is nat-

ural and the other part (reason) isn’t. Nor are we convinced by Davidson’s monist response, for we cannot see how events that are merely conceived as mental could have any causal impact in an exclusively physical universe. What we should learn from these failures, is to cling to Aristotle’s idea that a normal mature human being is a rational animal, without losing Kant’s idea that rationality operates freely in its own sphere. In other words, we mustn’t let the realm of reason come apart from the realm of nature, for the Aristotelian idea of a rational animal implies the notion of a nature in which reason – not natural law – is autonomous. This thought leads us to the next and possibly only remaining option, the one known as ‘property dualism’. This is the theory that human beings do indeed consist of a single (physical/material) substance, but one which has at least two irreducible types of properties.

Property Dualism and Emergent Properties

One way of seeing how a property dualist could respond to Kant’s challenge is to study Timothy O’Connor’s book *Persons and Causes* (2000), which argues that free will is realised through agent causation and is compatible with a substance monism. ‘Agent causation’ is a name for Kant’s thesis that a rational agent can only be conceived as free if we regard his actions as not determined by causes external or alien to himself. O’Connor’s two basic assumptions are (i) that top-down, mind-to-body (‘sink the *Bismarck*’) causation is *sui generis* – in other words, uniquely of its own kind and hence in an altogether different category to physical causation; and (ii) that this powerful new causal property is *natural*, because it emerged from and yet was not determined by the human body’s micro-elements and their complex interactions. In this case everything depends on the credibility of there actually being such things as ‘emergent properties’. So what are they?

An emergent property is something newly added to nature. The term distinguishes properties that may have emerged from those which gradually evolved. An evolved property is one produced by natural selection, a slow process that weeds out biologically harmful responses to the outside world and preserves and refines those which aid survival. The mechanism of natural selection obviously could not begin until a variety of organisms existed; so nature had already produced the variations before selection could start to act. Darwin knew this, but excluded as beyond his own enquiry the question of how the variations – indeed life itself – had originally emerged from matter. His work explained how organisms adapted to this variety, but not how the variety itself was created.

The key difference between evolution and emergence is that an organism evolves in response to the conditions outside of itself; whereas an emergent property arises from an organism’s own internal structure. Thus the property ‘mind’ is an emergent from life, and life itself is an emergent from matter. This would mean that mind is the cause, not the consequence, of brain development; and life is the cause, not the consequence, of the organisation of matter. So mind would determine the neural process (the physical element), and not be accidental to it – not the froth on the wave.

This sketch lets us see why O’Connor wants the macro-level properties and processes of the mind to be rooted in and

emerge from the micro-level structure of matter. Because if mind (as an emergent phenomenon) is grounded in matter it is *part of nature*; so the rational will, too, is natural – not non-natural, as Kant thought. The realm of reason does not after all come apart from the realm of nature.

A property is emergent when it introduces a qualitatively new macro-level feature into the world. A qualitatively new property is one that transcends or goes beyond the network of base properties which gave it birth. Mind exists as an autonomous property over and above a whole bunch of micro-physical goings-on. An emergent property is qualitatively new for the further reason that it brings in its train a causal power which is *sui generis* in nature with respect to its organic or physical source. The emergence of mind with its own power to cause the body to act ('reason's peculiar causality') brought about a fundamental change in the behaviour of the parent organism: *Homo* became *Homo sapiens* – a *rational* animal. In a human being the capacity to exert top-down causation is a power freely to choose one's course of action for *reasons*. Or as O'Connor puts it (in words remarkably close to Kant's), "Our reasons move us to act as we do, not as external pressures but as *our* reasons, as our own internal tendencies to act to satisfy certain desires or aims."

We have looked at two ways of responding to Kant's challenge to reconcile free will and natural necessity without having to go outside of nature as he himself did. We wanted a wholly natural solution, as did Davidson, one which assumes that human beings are material stuff through and through. In Davidson's anomalous monism all events are physical, so every mental event relies on a physical twin. Though both kinds of event spring from a physical source, Davidson's theory licenses our freedom because there are, apparently, no laws to determine how they interact. This struck us as too odd to accept. So we turned to property dualism, as deployed by O'Connor in his argument that mind is a *sui generis* property, with its own causality, that emerged from the complex stuff of which we're made. Property dualism such as this has the virtue of making our freedom and autonomy purely natural phenomena; just as consciousness – our awareness of the world and ourselves – presumably also emerged from nature. And eventually, it feels safe to assume, the twin gifts of language and logical thought flowed in their turn from the emergence of mind in *Homo sapiens*.

© STUART GREENSTREET 2012

Stuart Greenstreet, a business manager and writer by trade, began philosophy in the evenings at Birkbeck College, London, before graduating from the Open University, followed by further philosophy at the University of Sussex.