

A NATURALISTIC ACCOUNT OF FREE WILL

GONZALO MUNÉVAR

Introduction

Concerning the problem of the freedom of the will, Moritz Schlick once wrote that "... it is really one of the scandals of philosophy that again and again so much paper and printer's ink is devoted to this matter, to say nothing of the expenditure of thought, which could have been applied to more important problems ..."¹ It was particularly distressing to Schlick to have to take up this "pseudo-problem," which he thought to be a problem at all only because of a misunderstanding, as David Hume with exceptional clarity (presumably) had already demonstrated.

I must confess a great deal of sympathy with Schlick's impatience. Nevertheless, like Schlick, I find myself having to address it, in part because I find neither Hume's solution nor Schlick's reiteration of it satisfactory, and in part because it seems especially daunting to the evolutionary naturalism that I have advocated elsewhere.²

The problem is usually thought to be that we are not morally responsible if determinism is true, for if determinism is true we are not really free agents. The reason is that determinism implies that every action is causally necessitated, but if so, we can never act otherwise, and if we cannot act otherwise we are not free.

The problem seems particularly acute for naturalist points of view, since naturalism would treat persons as arrangements of matter, and thus it is ultimately consistent with some form of determinism. Naturalism would, then, imply that free will is an illusion and so is moral responsi-

¹ Schlick, M., "When is a Man Responsible?" reprinted in Enteman, W.F., *The Problem of Free Will*, Scribner's Sons, 1967, p. 184.

² See for example "Evolution and Justification," *The Monist*, Vol. 71, No. 3, 1988, pp. 339–357, and *Radical Knowledge*, Hackett, 1981.

bility. I think, however, that naturalism provides the most likely approach to the solution of the problem. This paper will thus aim to demonstrate that naturalism is compatible with the freedom of the will.

I shall begin by commenting on some recent naturalistic views on the problem. Although they fall short of the mark, they will prove to be helpful and instructive. I shall continue by discussing why indeterminism cannot account for freedom of action. I will then criticize some widely accepted solutions to the problem. And in the final section I will sketch a naturalist account of how the self determines the will, which is, I believe, the key to the solution.

Naturalist Views of the Problem

An intriguing example of a naturalist view can be found in Francis Crick's *The Astonishing Hypothesis*, in which he announces not only his theory of free will but also the location in the brain of the organ of free will.³ Two aspects of Crick's approach are important to my discussion. The first is his theory. Crick sets out to explain what philosophers would call the "phenomenology" of free will, that is, why it appears to us that we have freedom of action. Crick starts from the sensible assumption that some part of the brain "is concerned with making plans for future actions without necessarily carrying them out."⁴ We may, of course, be conscious of such plans. Crick further assumes, correctly I believe, that the actual working out of these plans (the "computations") are normally not open to our consciousness. We are aware only of the "decisions" taken by the brain, of the plans themselves. And finally, we are aware of the decision to act on one of these plans (e.g., to move), "but not of the computations that went into the decision."⁵ Thus even if the workings of the brain are completely deterministic, this feature of consciousness bars our "direct" access to them, and therefore we are aware only of decisions untangled by deterministic mechanisms. In other words, even if the deterministic mechanisms are there, we cannot be aware of them, and thus we have the experience of acting "freely." Moreover, some of these mechanisms may be deterministic but chaotic (this idea he attributes to Patricia Churchland), which would make their outcome seem unpredictable.

³ Crick, F., *The Astonishing Hypothesis: The Scientific Search for the Soul*, Scribner's Sons, 1994, pp. 265–268.

⁴ *Ibid.*, p. 266.

⁵ *Ibid.*

The role of consciousness is much reduced in Crick's account (relative to the extraordinary significance philosophers characteristically assign to it). A brain so described, Crick believes,

can attempt to explain to itself why it made a certain choice (by using introspection). Sometimes it may reach the correct conclusion. At other times it will either not know or, more likely, will confabulate, because it has no conscious knowledge of the "reason" for the choice.⁶

This result is consistent with the work by Crick and others on the relationship between consciousness and other functions of the brain to the effect that conscious deliberation is not always necessary for rational decision—let alone, as in this case, for mere decision, whether rational or not.⁷

Crick's hypothesis seems to lead to the conclusion that free will is an illusion. Others have arrived at this conclusion before, but not many have offered as compelling an explanation of *how* that illusion comes about. It seems to me that Crick's approach has one cardinal virtue and one cardinal flaw. Its virtue, which is a naturalistic virtue, is that it places some empirical constraints on our thinking about free will. Crick assumes that whatever free will turns out to be, it is a faculty of the brain, and thus studying the brain may reveal some of its important features. As soon as we take this approach, we realize, for example, the possibility that free will, as any other faculty of the brain, may malfunction; we realize, that is, that some people may have a defective, or non-existent,

⁶ *Ibid.*

⁷ Until rather recently scientists tended to place the same importance on reflection. As late as 1973, Salvador Luria wrote that "Human behavior is conscious behavior and by virtue of that fact man is more than another animal." (*Life: The Unfinished Experiment*, Scribner's Sons, 1973, p. 146.) Darwin himself thought that the moral sense comes from the application of the higher mental powers to the social instincts, and by those powers he meant memory, anticipation, and the power of reflection (cf. Bradie, M., Note 41). It is true, of course, that deliberation often amounts to conscious deliberation, but it need not be so. Think of the hundreds of little, and sometimes big, decisions that we make in the course of the day: while driving, walking, taking the stairs, dancing, painting, or playing a game. Some may be the results of habit, automatic "subroutines" of the brain, but in many cases we have to own up to them. In sport, for example, good decisions often depend on our reading the situation correctly and quickly and making a "split-second" choice appropriate to the situation. In war those quick decisions could not have more serious consequences. It would be strange to disown those decisions, to say that we did not want to act that way, or that we are not responsible for them but will accept responsibility only for those actions about which we deliberated at length.

“organ” of free will. And indeed there appear to be such people, that is, people who seem not just unable but actually uninterested in choosing courses of action. Moreover, this handicap can be attributed to a lesion to a part of the brain that seems involved in making plans (the anterior cingulate sulcus, a region near the top and towards the front of the brain, which receives “many inputs from the higher sensory regions and [is] near the higher levels of the motor system”). As we will see below, this sort of hypothesis can be helpful in unraveling some of the puzzles of free will.⁸ Nevertheless, Crick’s explanation is flawed in that it leaves the main philosophical problem untouched. If we are not free agents, how are we then responsible for our actions?

Naturalists have tried several approaches to get around this problem. I will discuss some important ones. A rather common approach among scientists (as opposed to professional philosophers) is to analyze determinism in practical terms. E.O. Wilson, for example, describes the many cognitive abilities of a honeybee: it has memory, knows the time of the day, learns the location and quality of several flower fields, and responds vigorously and “erratically” to physical challenges. The bee thus

appears to be a free agent to the uninformed human observer, but again if we were to concentrate all we know about the physical properties of thimble-sized objects, the nervous systems of insects, the behavioral peculiarities of honeybees, and the personal history of this particular bee, and if the most advanced techniques were again brought to bear, we might predict the flight path of the bee with an accuracy that exceeds pure chance.⁹

⁸ A result of this line of thought is that the question of free will may extend into the animal kingdom well beyond humans (cf. Wilson’s discussion of honeybees below), for it may be difficult to draw a sharp line of demarcation, as is the case with other traits that have resulted from natural selection, including intelligence (which does not require, by the way, that mice, say, ask themselves whether they are intelligent). A similar approach would best fit the emergence of consciousness. Some philosophers may think that language is necessary for conscious reflection, but Paul Churchland has made an excellent case for the claim that “[T]he cognitive priority of the preverbal over the verbal shows itself, upon examination, to be a feature of almost all of our cognitive categories.” (*The Engine of Reason, the Seat of the Soul*, MIT Press, 1995, p. 144.) An advantage of naturalism is the possibility of future comparative animal studies that would permit us to understand better human decision making by seeing how it differs from animal decision making.

⁹ Wilson, E.O., *On Human Nature*, Harvard University Press, 1978, pp. 72–73.

The point is that to human observers using such techniques the future of the bee is determined to *some* extent, but "in her own 'mind' the bee, who is isolated permanently from such human knowledge, will always have free will."¹⁰

In the case of human beings, we are so much more complex than bees that only techniques beyond our present imagining could hope to achieve even the short-term prediction of the detailed behavior of an individual human being, and such an accomplishment might be beyond the capacity of any conceivable intelligence. There are hundreds of thousands of variables to consider, and minute degrees of imprecision in any of them might easily be magnified to alter the action of part or all of the mind.¹¹

Wilson suggests also an analog of Heisenberg's uncertainty principle, in which observations of human behavior alter that behavior. I suspect that he incurs here in an analog of the mistake of claiming that we could never discover the nature of life because in studying it we alter it. In any event, all these reasons lead Wilson to suppose that no nervous systems may gain enough knowledge of the mind to "know their own future, capture fate, and in this sense eliminate free will."¹² You and I are consequently free and responsible in the fundamental sense that the "detailed histories of individual human beings [cannot] be predicted ... by the individuals affected or by other human beings."¹³

This is not to say, of course, that we cannot predict general tendencies in behavior, or in fact what a particular person is likely to do. Wilson's concern is the ability to predict beyond statistical regularities. The scientific aspect of Wilson's analysis has been strengthened in the subsequent twenty years of great success in the neurosciences. As Paul Churchland explains, the brain is a non-linear system, that is, a system

in which, at least occasionally, even the tiniest of differences in its current state will quickly be magnified into very large differences in its subsequent state. Since we can never have *infinitely* accurate information about the current state of any physical system, let alone a system of the complexity of a living brain, we are doomed to be forever limited in what we can predict about such a system's unfolding behavior, even if

¹⁰ *Ibid.*, p. 73.

¹¹ *Ibid.*

¹² *Ibid.*, p. 74.

¹³ *Ibid.*, p. 77.

there are, and even if we happen to know, the inviolable laws that govern the system's behavior.¹⁴

Wilson's practical analysis, then, equates determinism with predictability and shows how perfect predictability is practically impossible. The first part of his analysis is by no means unreasonable. One of the oldest forms of the Problem of Free Will was precisely that of predestination: If God already knows how we are going to choose, in which sense can we be said to be free? A similar metaphysics of free will would be consistent with the views of those philosophers who take to heart Einstein's notion that time is an illusion. In the theory of General Relativity all events are described by four coordinates, time being one of them. Time becomes an objective coordinate whether we phenomenologically describe it as past, present or future. We might thus say that time is already laid out: Nothing can be except what is. For all we care, thus, our future actions are as objective as our past actions. We cannot deviate from what we are bound to do any more than are able to change what we have already done. I suspect that this view of time is done in by quantum physics (as is determinism itself, see below), but be that as it may, it is clear that some connection seems to exist between determinism and predictability. Wittgenstein, for example, claimed that freedom implies ignorance of what we are going to do (prior to deliberation, that is).¹⁵

Nevertheless determinism and predictability are not the same thing. As Churchland points out in the continuation of the passage quoted above, "Such systems are *strictly deterministic*, in the sense of being law governed, but they are nevertheless unpredictable, beyond their statistical regularities, by any cognitive system within the same physical universe."¹⁶(My emphasis.) The problem of the freedom of the will involves determinism, not prediction. The reason is that, as William James saw, freedom requires alternative possibilities, but for determinism only what does happen is possible. But then, "... what sense can there be in condemning ourselves for taking the wrong way ... unless the right way was open to us as well?"¹⁷ It is for this reason that Wilson's suggestion ultimately fails. For in Wilson's practical freedom of the will our actions are

¹⁴ Churchland, P.M., *op. cit.*, p. 113.

¹⁵ See the discussion by Paul Horwich in his *Asymmetries in Time*, MIT Press, 1987, p. 204.

¹⁶ Churchland, P.M., *op. cit.*

¹⁷ James, W., "The Dilemma of Determinism," reprinted in Enteman, *op. cit.*, p. 69.

still determined, whether we can predict them or not, and thus when we take the wrong way we do not really have the right way open to us.

There are those who would wash their hands off the whole problem and settle for some sort of practical "freedom." Churchland, for example, insists that the complexity of the brain (which creates the unpredictability) does provide us with the capacity for genuine spontaneous activity, for endless variety in our behavior, and this capacity is very important, even though

It would be foolish to mistake such (genuine) unpredictability for what philosophers and theologians have often hoped for in the way of free will. That term was typically meant to apply to a human capacity that *transcended* the causal order, whereas the dynamical picture [of the brain] keeps us firmly embedded within the causal order.¹⁸

As sensible as these remarks sound, they would be philosophically more reassuring if we had an account of moral responsibility to go with them. Or if we could undermine the connection between determinism and moral amnesty. I will concentrate on the second of these options in the next two sections.

Indeterminism and Free Will

At first sight no discovery would seem to be as important for the issue at hand as the discovery that determinism is false. Thus we need to understand why the extraordinary success of quantum physics has not settled the matter. After all, quantum physics, at least in its orthodox interpretation, tells us that at the most fundamental level nature is to be seen as probabilistic, not because of ignorance (which is the traditional interpretation of probability in science) but because it is a basic property of physics. This is not to say that there are no deterministic processes or laws in the universe, but rather to deny that all phenomena are determined. This result raises the hope that the mind is not wholly determined, and that perhaps there is room for free will after all. As we will see in this section, however, this hope is misguided.

There are two main ways in which quantum indeterminism is introduced into the dispute. The first would have the (non-material) mind somehow interact with the brain at the quantum level. In this sense quantum physics would come to the aid of metaphysical (or substance)

¹⁸ Churchland, P.M., *op. cit.*

dualism, i.e., the view that mind and matter belong to separate "realities" or perhaps "dimensions." I presume this is what Churchland had in mind (no pun intended) when he spoke derisively of a human capacity that transcended the causal order. Dualism can take on three forms. First, the mind and its physical correlate (whether the brain or the whole central nervous system) exist in parallel dimensions: whatever happens in the mind has its counterpart in the brain. Second, the mind is an epiphenomenon of the brain, something like an innocuous halo on the causal order. Third, which was Descartes' favorite, the mind and the brain interact, viz., the mind interferes in the causal order.

Given the great recent success of the neurosciences in understanding the mind, the first two forms of dualism have little explanatory power and seem hopelessly *ad hoc*. In any event, neither epiphenomenalism nor parallelism can help with our problem of free will. Epiphenomenalism depends entirely on the causal (or not) order of the brain and thus neither adds nor subtracts anything to the issue of freedom. As for parallelism, if I raise my hand presumably of my own volition, but my action turns out to be determined, its counterpart in the mental dimension will also be determined. The perfect synchronization of the mental and physical dimensions will make it so that the circumstances that determine my brain processes (or else their counterparts in the mental realm) also determine my equivalent mental processes. This leaves us with the third option: interactionism, a view already saddled with having to overcome a great many implausibilities, chief among them the violation of the laws of nature (e.g., the conservation-of-energy law). The idea is, however, that quantum phenomena could come to the rescue here, for at the level of the extremely small perhaps the mind could guide the way these otherwise indeterminate events go. Once the mind has done its micro-job, the brain would amplify the quantum events and lead to action. Although I realize that the mind still is very mysterious, and that quantum physics also seems mysterious, I do not share the hope that they must thereby be connected. For in this connection we violate exactly the conditions that make quantum physics the paradigm of indeterminism. If mental events determine quantum events, if the latter are no longer fundamentally probabilistic, then quantum physics is false and we have not helped the cause of free will at all.¹⁹

¹⁹ Another possibility would be to have the non-physical mind somehow collapse the wave packet. I have always found difficult to understand suggestions to that effect.

These considerations show, incidentally, that we do not place ourselves in a better position to solve the problem of free will merely by eschewing naturalism and taking a higher, spiritual road instead.

A more naturalistic approach is to argue that brain processes have their origin in quantum events (which are amplified) and that therefore the brain is ultimately indeterministic. Unfortunately there are serious problems with this suggestion. One of them is that the acceptance of the suggestion would force us to accept also the notion that there are no deterministic processes whatsoever, for all physical processes can be said to have their origin in quantum events (e.g., when a billiard ball hits another, the electromagnetic fields of their orbital electrons first come into contact). Indeed, quantum theory draws a classical limit to which the statistical quantum properties tend, which means that outside of microphysics nature still is deterministic. Another problem, which I will not discuss here, is that the suggestions that have actually been put forward suffer from severe scientific implausibility as soon as we take into account not just physics but neurobiology (for a vivid example, see the devastating critique by Rick Grush and Patricia Churchland of Roger Penrose's attempt to link quantum physics and the brain via the microtubules of neurons).²⁰

Ilya Prigogine has argued that complex systems, particularly those that are far from equilibrium, can be indeterministic. Even molecules in a real (non-idealized) gas respond to non-local and atemporal resonances, he claims, if nothing else because of their persistent collisions, which means that their motions depend on properties of the system.²¹ This is not a plain emergent property of the system (i.e., that cannot be determined by knowing the properties of their constituents), it seems to me, but a stronger case in which the properties of the constituents are affected by the properties of the system. I am sure that Prigogine's interpretation of a system of gas molecules is open to dispute; what matters here, though, is whether the brain exhibits this strong property of emergence, and it seems to me that it does. Indeed, if anything the brain presents a clearer example, as we will see below. In systems far from equilibrium at least, Prigogine holds, we can in principle predict only statistical properties, and thus probabilities become a fundamental property of

²⁰ Grush, R., and Churchland, P.S., "Gaps in Penrose's Toilings," *Journal of Consciousness Studies*, 2, No. 1, 1995, pp. 10–29.

²¹ Prigogine, I., *The End of Certainty: Time, Chaos, and the New Laws of Nature*, Free Press, 1997.

nature. I suspect that in developing this indeterminism born from instability and chaos, Prigogine has fallen into a subtle confusion between predictability and determinism, but this is not the place to do justice to his rather complicated analysis.²²

Nevertheless, consider that the brain has about 100 trillion modifiable synaptic connections, that each of these connections can assume a large number of synaptic weights, and that these weights depend not only on the sensory signal arriving at the synapse of the neuron, say, and the structure of the neuron, but on the effects of other neurons on it. The sources of these effects, furthermore, are not restricted to the hundreds, or even thousands of neurons around it, but may be found also in neurons located in remote parts of the brain that connect in feedback loops to the neuron in question (indeed, as Paul Churchland points out, often what he calls recurrent pathways are more numerous than those carrying information forward from the senses). The result is that the brain systems or networks achieve temporary states of stability by "cycling through" many tentative synaptic weights and adjusting them so as to achieve a suitable accommodation to the goals of the moment. Thus a neural state is emergent in the sense that the weights of the synaptic connections that constitute it are not sufficient to determine it, and also emergent in the sense that those weights are also partially dependent on the neural state itself. But even though neural networks seem to fulfill Prigogine's criteria, every one of the relevant interactions obeys deterministic laws and, as the previous quotation from Churchland indicates, such networks should thus be considered to operate within the causal order. The brain is, then, both unpredictable and deterministic.

If perchance our actions were the result of indeterminism, of chance, the supporters of free will would face a most unwelcome discovery: Chance is also incompatible with free will. As Reid, and later Hume, pointed out, if our actions are the result of chance we have no control over them, they are thus not truly *ours*. In Gary Watson's words, "What destroys freedom ... is the lack of self-determination and that results both when the will is determined by other events or states of affairs and

²² Prigogine, following Poincaré, argues that the atemporal resonances make it impossible to integrate over the trajectories of the gas molecules. Presuming that this is right, we still need to consider whether the non-integrability amounts to in-principle unpredictability. But even presuming that it does, from non-integrability we cannot conclude indeterminism, nor can we do so from unpredictability alone, as we have seen.

when it is not determined at all."²³ Watson then puts his finger on the true nature of the problem: "The negative requirement that the will not be causally necessitated by antecedent events is dictated by the positive requirement that the will be determined by the self."²⁴

Naturalists, scientists in particular, have been looking in the wrong place, their quest has been distorted by their emphasis on only one or two aspects of the problem: predictability, determinism, or both. Our task should instead be to provide a naturalistic account of how the self determines the will. But before this task is undertaken more directly, one more hurdle must be overcome: the claim made by some philosophers that there is really no problem of free will at all. Otherwise we would spend even more time and effort on a task for which there is no need.

Philosophical Solutions

Many philosophers of science tend to ignore the problem of free will because they think that Hume already solved it a long time ago. I disagree with them. Hume's first move was to undermine the notion of Necessity (the inevitability that James understood was part and parcel of determinism). Our notion of necessity and causation, says Hume, "arises entirely from the uniformity observable in the operations of nature, where similar objects are constantly conjoined together, and the mind is determined by custom to infer the one from the appearance of the other."²⁵ We are thus wrong when we believe we "perceive something like a necessary connection between the cause and the effect."²⁶ This is very important, for the regularity of human behavior is crucial to our social lives. Indeed, Hume asks, "Where would be the foundation of *morals*, if particular characters had no certain or determinate power to produce particular sentiments, and if these sentiments had no constant operation on actions?"²⁷ Thus we are wrong when we suppose that a difference exists between "the effects which result from material force" and

²³ Watson, G., "Free Will," in Sosa, E., and Kim, J., eds., *A Companion to Metaphysics*, Basil Blackwell, 1994, p. 178.

²⁴ *Ibid.*

²⁵ Hume, D., "Of Liberty and Necessity," from *An Enquiry Concerning Human Understanding*, Open Court Publishing Co., 1907. Passages quoted in this essay come from Section VIII, Parts I and II, reprinted in Enteman, W.F., *op. cit.* This quote comes from p. 166.

²⁶ *Ibid.*, p. 174.

²⁷ *Ibid.*, p. 172.

those "which arise from thought and intelligence."²⁸ As Hume puts it, "this experimental inference and reasoning concerning the actions of others enters so much into human life that no man, while awake, is ever a moment without employing it."²⁹ But in both cases "we know nothing further of causation of any kind other than merely the *constant conjunction* of objects and the consequent *inference* of the mind from one to another."³⁰

A solution cannot be found, Hume argued, "as long as we will rashly suppose that we have some farther idea of necessity and causation in the operations of external objects."³¹ while in the case of human action "we feel no such [necessary] connexion."³² Once we realize that the same notion of "necessity" operates in both, and that there is nothing really necessary about causation, the way is open to a solution. And the solution is simple: when we think carefully about cause and effect we come up with an innocuous connection that does not conflict with freedom of action. Such a conflict comes about because of our rash judgment about causation and because of one additional misunderstanding: our opposing liberty to necessity and not to constraint. Hume follows Hobbes in claiming that by liberty, when applied to voluntary actions, "we can only mean *a power of acting or not acting according to the determinations of the will*; that is, if we choose to remain at rest, we may; if we choose to move, we also may."³³ The question is settled, then, for "this hypothetical liberty is universally allowed to belong to every one who is not a prisoner and in chains."³⁴

Hume's solution has been adopted by a variety of philosophers. Mill, for example, tells us that "Even if the reason repudiates, the imagination retains the feeling of some more intimate connexion, of some peculiar tie, or mysterious constraint exercised by the [cause] over the [effect]."³⁵ It is precisely in that mysterious constraint that the conflict arises, for "We

²⁸ *Ibid.*, p. 174.

²⁹ *Ibid.*, p. 172.

³⁰ *Ibid.*, p. 174.

³¹ *Ibid.*, p. 175.

³² *Ibid.*, p. 174.

³³ *Ibid.*, p. 176.

³⁴ *Ibid.*

³⁵ Mill, J.S., "Of Liberty and Necessity," from *A System of Logic Ratiocinative and Inductive*, Longmans, Green, Reader and Dyer, 1872, II, reprinted in Enteman, W.F., *op. cit.*, p. 257.

know that we are not compelled, as by a magical spell, to obey any particular motive."³⁶ Schlick calls our attention to the distinction between natural law and the law of the state. The first one involves determinism and its "necessity" consists merely in its being universally valid. The second involves compulsion. The problem of free will stems from the confusion of the two, that is, from ascribing to natural law the compulsion characteristic of a law of the state. For him, as for Mill and Hume, and many more recent philosophers, the entire controversy can be avoided if we only pay proper attention to the meaning of words. And it is in that proper attention that we understand that causality, far from being incompatible with responsibility, is required by it. "We can speak of motives only in a causal context," Schlick says, "thus it becomes clear how very much the concept of responsibility rests upon that of causation, that is, upon the regularity of volitional decisions."³⁷ Hume's very point.

Nevertheless, this approach to the problem is found wanting when we remember William James' simple observation: we cannot be blamed for taking the wrong way when the right way was not open to us. Let us look first at the much-debated Humean account of causality. Suppose that a wrecking ball is suspended ten meters over a delicate porcelain vase. If we let go of the ball it will pulverize the vase. Any deviation from this result will require special circumstances (e.g., a hidden gigantic electromagnet). But if all the relevant circumstances are accounted for, and they are as first described, when the ball is set loose it *will* destroy the vase. It *will* happen that way, to the exclusion of alternatives. Hume himself would not deny this point. We thus need no belief in mysterious constraints or magical spells to realize that determinism leaves no options open. Second, we may realize that the problem does not arise from confusing natural law and the law of the state, that is, from attributing compulsion to natural law (or causality). On the contrary, compulsory law may give us an excuse, or a justification, for not acting otherwise, but we seldom, if ever, attribute to it a universal character. Sometimes we disobey it, and there are laws that most people disobey (e.g., speed laws for automobiles). The problem with acting in accordance with natural law (which we cannot avoid) is precisely that it places us in the same category as the wrecking ball, regardless of how special and free we feel. Einstein once wrote that

³⁶ *Ibid.*

³⁷ Schlick, M., *op. cit.*, p. 192.

If the moon, in the act of completing its eternal path round the earth, were gifted with self-consciousness, it would feel thoroughly convinced that it would travel its path on its own, in accordance with a resolution taken once and for all.³⁸

Compulsion may exculpate us, but it allows us to remain moral agents. If a man with a gun to his head is told to kill another, and complies, we excuse him. But if he is thrown from a tall tower and thus kills another on the ground, the question of excusing him does not come up, for being a falling body is completely out of the moral sphere. Therefore Hume's, Mill's, and Schlick's "proper attention" to the meaning of words does nothing to solve the problem of free will. The problem is created by determinism precisely because it seems to put our actions on a par with the behavior of falling bodies —precisely because when we act "wrongly" we do not seem to have the "right way" open to us.

Nor is further "linguistic" analysis of the "logic" of freedom and responsibility likely to improve matters much. The reason is simply that such analysis belongs ultimately to the "phenomenology" of the issue, even if limited to the level of discourse, and appeals to such phenomenology are rendered moot at best, and question-begging at worst, in a dispute in which the challenge is to show that it is not an illusion. Imagine an extremely sophisticated android who, unbeknownst to him, receives radio instructions from me as to how to behave (a suitably wired human would do also). In accordance with Crick's account of the brain, the android is aware only of the "decisions" (stand, sit, etc.) and not of the fine work of his brain's mechanisms (Crick's "calculations"). Unlike my brain, however, his brain is affected by my radio signals and the decisions it arrives at are chosen by me. If I instruct him to remain at rest, he may; if I instruct him to move, he may. That is, nothing prevents him from carrying out the action that his consciousness informs him is his brain's decision —his will— for he "is not a prisoner and in chains," as Hume said. Or to use Hobbes' notions, once he "wills" something he has the power to do it. The android is then a free agent because "he can do if he will, and forbear if he will."³⁹ The android may thus satisfy what Hobbes, Hume, Mill, and Schlick thought we "mean" by freedom of the will, but he is not really a free agent because his will is determined by me.

³⁸ Quoted in Prigogine, I., *op. cit.*, p. 13.

³⁹ Quoted in Watson, G., *op. cit.*, p. 176.

Other conceptual and linguistic analysis may perhaps uncover different important "meanings" concerning freedom and responsibility, but we will not make much headway with our problem as long as the much despised "free-will metaphysician" can claim that our will is determined by something other than its own self.

This point extends to another "discovery" of linguistic philosophy, namely the claim that there is no problem of free will because the language of determinism belongs properly to material things, whereas the language of thought, which presumably covers such notions as freedom and responsibility, involves not causes but reasons and intentions. But as Gary Watson argues, the traditional problem is not avoided in this way.

For if physical determinism is true, it is impossible for your body to move in any way other than its actual motion. This means that it is not possible that you will *move* your body in any different way, and hence that you will act in any way that requires a different bodily motion.⁴⁰

Moving the discussion to the realm of reasons and intentions still does not explain how we can act otherwise. As Watson continues, "If it is physically determined that your arm does not go up during a certain period, then it is not possible that you will signal the waiter, say, by raising your arm." Unless, of course, this conceptual dualism, as we may call the philosophical view under discussion, manages to explain "how it can be the case that you are able to raise your arm during a time when it is causally impossible for your arm to go up."⁴¹

Naturalism and Self-Determination

The question is whether the will can be determined by the self. I believe that it can, that it is typically so determined, and that we can understand how once we adopt a naturalistic approach. In the rest of this section I will sketch such an approach, with enough detail, I hope, to constitute a good first approximation.

Let us begin with a naturalistic account of the self. Most, if not all, living things have at the very least a primitive sense of self. A single-cell organism, for example, is organized to distinguish what belongs to it from what does not. Indeed, most bacteria will identify invading cells and use a variety of chemical means to destroy those invaders. In a

⁴⁰ *Ibid.*, p. 180.

⁴¹ *Ibid.*

primitive sense we may say that organisms can tell self from non-self. As the complexity of the organism increases, a nervous system may develop to coordinate the different organs and functions. However loose this organization may be in some "primitive" animals, it is more than a mere black box by means of which the inputs to sensory receptors are transformed into behavioral outputs. The behavioral response, if any, is modulated by the animal's sense of itself and of its relation to the world, a rather complex achievement that depends on internal information (gathered through proprioceptors) and its mutual fine-tuning with the external sensory information. As the complexity of the nervous system increases, so does the number of possible ways of coordinating and modulating neural systems. Visual perception, for example, depends on feedback loops from the other senses, including the internal senses, as well as from modalities that involve memory, expectation and even emotion. This increase in complexity leads to an increase in plasticity and to the appearance of what we might call intelligence. In social animals presumably the brain is genetically "wired" if not for cooperation, for the capacity to cooperate, as Darwin argued over a century ago.⁴² This disposition toward cooperation, made up of our "social instincts," leads to the origin of morality in intelligent creatures. In a complex brain, of course, the moral emotions that may result from such a disposition are subject to a great amount of plasticity. It is from those emotions, in any event, that, according to Darwin, we derive our moral conscience. In this connection, with the development of moral responsibility in the brain, we should find the relevant features of free will. After all, the problem of free will revolves around the question of whether the decisions we take concerning those actions thought to fall within the sphere of moral responsibility are determined by the self or not.

In this naturalistic account the self is, of course, none other than the brain. So the question is whether the brain determines the will, i.e., whether it has the power to make those decisions open to moral evaluation. The answer is that the brain does. Lest we think that this answer is too facile, let us realize that the brain's self-organization pits the organism as an independent entity vis-a-vis the world. When discussing Prigogine's claims against determinism, which he based partly on his analysis of the order exhibited by many states far from equilibrium, I

⁴² See the discussion by Michael Bradie in his *The Secret Chain: Evolution and Ethics*, Suny Press, 1994, pp. 58-64.

brought up the notion of strong emergence and explained how it applied to the brain. This strong emergence goes beyond complexity and plasticity in establishing the operation of any particular brain as *sui generis*, since any small variation from brain to brain—and there are large variations—can be greatly amplified, as Churchland pointed out above, and since a brain will work through a given situation in *its own* mysterious ways. The question here is not practical, or even in-principle, unpredictability, as it was for Prigogine, but self-determination.

Although the world exerts an influence on the brain's decisions, either as present stimuli, or as past experience of the species (which gives us our basic modes of thought and perhaps our basic moral emotions), a strongly emergent system such as a human brain amounts to a pocket of the world ruled by emergent "laws" of its own. There may be a worry about whether we may speak of universal laws in the case of the brain with any greater assurance that we can in the case of evolutionary biology, but there is no question that the brain exhibits emergent causal relationships and that its causal system is generally self-sustaining and independent. Again, by independent I mean simply that there is a discontinuity between the rest of the world, its natural laws included, and the new emergent "laws" by which each individual brain interprets a situation, finds it relevant, evaluates it, and finally decides how to deal with it. Natural laws, of course, operate in the brain, with the many elements that come to play a role in any decision organized in a manner that roughly resembles that of other brains, but that also depends on the peculiar characteristics of each brain. It is that organization that places the whole of those elements beyond the behavior of mere falling bodies, just as it is the organization of the elements of the brain that makes their joint action intelligent. That is how intelligence arises out of matter: as a peculiar systemic property of a peculiar dynamic organization of that matter. That is indeed how mind, the self, arises out of matter. We may have this image of the little cogs and wheels of the brain blindly turning inside our brains in obedience to indifferent and universal laws, but this image ignores that these very cogs and wheels are not external factors impinging on the self but internal elements of a complex whole which is, to borrow the cliché, more than the sum of its parts. Moreover, not only are these cogs and wheels elements of the self, but their very character as elements is determined by the complex whole to which they belong. Neurons, for example, have their synaptic weights modulated by a complex array of influences from other neurons, as Paul Churchland has ex-

plained in his account of recurrent pathways in the brain.⁴³ It is reasonable to suppose that complex arrays of information higher in the brain's hierarchical structure (what I take to be Wilson's "schemata") interact, combine, and compete with other such complexes for central stage. The contribution of an element of the self to a decision then depends partly on the systemic influences of the self on that element.

How is this a solution to the problem of free will? To borrow Watson's words once again, "the question is how a series of natural processes (for which you are not accountable) can result in processes and events over which you do have control (for which you are accountable)."⁴⁴ Let us, then, see what sorts of processes and events we do not have control over and how they result in some over which we do control (and for which we are thus accountable). Schlick fairly ascribed to the free-will metaphysician the view that the will is determined by character and motives, and therefore we can do nothing about the way our decisions go, for we have no power over either: "the motives come from without, and my character is the necessary product of the innate tendencies and the external influences which have been effective during my lifetime."⁴⁵ Many motives, however, are internally generated—think of ambition or self-loathing, for example. And those which we may regard as external do not come entirely from without either. A hostile look from a stranger motivates me to search quickly for a heavy object that I can hold in my hand. But it does so only because I interpret his look as hostile, and because I read the situation as dangerous since there are no other people around, my foot injury will keep me from running, and so on. Something does not become a motive unless it is first interpreted by the brain in a particular way and then seen as relevant to a goal that the brain has decided to pursue. External influences do exist, of course. Almost everyone in the world would consider a pistol pointed at his head a motive for action, but even then the pistol has to be recognized as such (not a toy pistol, for example), other clues have to be read as corroborating the impression of danger in the situation, etc. However we interpret the situation, though, we may have no control at all over the event of having the pistol pointed in our direction. Once the event is processed and integrated into our mind it provides a motive and we may

⁴³ Churchland, P.M., *op. cit.*, pp. 97–150.

⁴⁴ Watson, G., *op. cit.*, p. 181.

⁴⁵ Schlick, M., *op. cit.*, p. 186.

make a decision about it. But this decision is guided by the characteristics of the particular brain, by its needs, desires, beliefs about past experiences (which would include true and false memories), by its emotions, and by the urgency of other decisions taken about the same time. Every decision is thus stamped with the very personal seal of the brain that has control over it. Some people would run, others would shoot first if armed, and still others would take the bullet to protect a loved one.

As for character, the pertinent external influences would be handled in the same way just described. The innate tendencies can indeed be strong: evolution and embryology are certain to forge in the brain a variety of dispositions, but the brain is also distinguished by its ability to adapt, to change, to learn, and thus to transform itself. Moreover, this transformation is to some degree a matter of choice. In this, character is not an exception. As Mill pointed out, a man's "character is formed by his circumstances ... but his own desire to mould it in a particular way is one of those circumstances, and by no means the least influential."⁴⁶ Indeed, Mill suggests that the feeling "of our being able to modify our own character *if we wish* is itself the feeling of moral freedom which we are conscious of." It seems to me that Mill is right at least to this extent: we do blame ourselves, at least in moments of candor, for bad character traits that we realize we could have overcome through firmer resolve as we formed our habits. Many internal processes may exert great influence not only on the formation of character, but on the decision making itself. But we should keep in mind that those processes, too, do not become factors in a decision until they are assimilated into the whole by emergent mechanisms. When the self determines a decision, it does so *qua* self, its choices are not forced upon it by factors over which it has no control. On the contrary, the self *qua* self is what controls all these factors, assigns them values within the system, makes them relevant, compares and combines them with other factors. Otherwise they would play no role in the decisions the self makes.

It is precisely when factors outside this organic assimilation and control by the brain determine a decision that we can correctly claim lack of moral responsibility. Let us apply the previous android test. If Peter (Peter's self) would have decided to stand up, but through radio signals I alter the decision so as to stay seated instead, Peter is not acting freely. It was not his brain that made the choice, even though the brain mecha-

⁴⁶ Mill, J.S., *op cit.*, p. 260.

nisms may be such that he cannot help but think that he made the decision himself. Likewise with drugs, brain injury, or any disease that alters the normal decision-making operations of the brain. When a disruption in the proper rates of neuro-transmitters render Mary completely unable to interpret a situation as she would under normal circumstances, or gives extraordinary significance to an event that would not be normally that important to her (as in drug-induced paranoia), we should exempt (not merely excuse) her from moral blame to the degree of her inability. Of course, we may blame her for the choice to take the drug, and therefore charge her with negligence, but the reason for this harsh judgment is that her brain was working normally then. And when Alzheimer's robs John from access to his past, when the continuity of his self is thus disrupted, we are again not entitled to assign blame.

When a man is himself, his decisions will not proceed from character and external influences with necessity. They will result instead from the integration of internal and external influences into an individual-specific, and strongly emergent causal system: his self. Naturalism thus allows us to conclude with confidence that the will is determined by the self.

Unlike Hume's, Mill's, and Schlick's notions of freedom, all of which fail the android, drugs, injury, and disease tests, this naturalist account of the will (the self's —the brain's— power to make decisions) does a better job of explaining the phenomenology of freedom and responsibility. It does so while explaining how events and processes over which we have no control are assimilated by the self (the brain) and as a result we engender processes over which *we* do have control and for which we are, therefore, morally responsible.

University of California, Irvine