

Pierre Jacob, *What Minds Can Do*

Cambridge: Cambridge University Press 1997, pp.299+xii

Among the many things minds can do, two general kinds of thing have inspired much of the debate in recent philosophy of mind. The first is minds' power to represent the world, their intentionality. The second is their power to generate action. The first power has seemed problematic to physicalist or naturalist philosophers, since it is hard to understand how a mere physical object—which is what a mind is—can 'direct' itself on things in the way characteristic of intentionality. The second power has seemed problematic not least because it is hard to understand how intentional states can bring about changes in the physical world without adding new mysterious forces into the world.

Pierre Jacob's book is an elegant, informative and ambitious treatment of these two problems. The first half of the book is a discussion of the first problem—how to 'naturalise' intentionality. Here Jacob adopts an 'informational' approach: minds carry information about their environment, (in the precise sense of the mathematical theory of communication). In response to the objection that this makes *misrepresentation* impossible, Jacob appeals to the idea that states of mind inherit their capacity to represent from their biological functions, in a way that is now familiar in the literature in recent philosophy of mind. The second half of the book distinguishes various problems of mental causation, and attempts to solve the different problems in different ways, as we shall see. In the middle, there are nice discussions of holism and the computational view of mind.

Jacob's approach is chiefly inspired by Fred Dretske and Jerry Fodor. From Dretske, Jacob takes the teleological solution to the misrepresentation problem, the 'component' view of behaviour and the distinction between structuring and triggering causes. From Fodor he takes the computational view of the causal basis of intentionality. And in the background is a non-reductive version of physicalism inspired by Donald Davidson.

One might wonder whether all these views are compatible. For, on the face of it at least, the computational theory of mind requires the supervenience of the mind on the brain, but informational teleosemantics rejects supervenience; the computational theory of mind seems committed to the inefficacy of semantic or intentional properties (there are no ‘semantic engines’) but Dretske and Jacob are committed to their efficacy; the teleological theory emphasises the biological basis of representation, but the computational theory emphasises abstract computational capacities which in principle could be had by a system which had not evolved at all; and it is very hard to see how the mechanistic reductionism of Fodor’s and Dretske’s approaches to intentionality can be squared with Davidson’s denial of psychophysical laws and his hermeneutic, holistic conception of the intentional. Of course, many would see some of these conflicts as only apparent. But my reading of Jacob’s attempts to find common ground between the views of these thinkers has tended to reinforce my superficial impressions.

Jacob’s work on the naturalisation of intentionality seems to me the most successful part of the book. As noted above, he appeals to Dretske’s informational teleosemantics: states of minds are natural indicators of—that is, they causally co-vary with—the phenomena that it is their biological function to represent. At first sight it might seem puzzling that the second problem (the causal efficacy of intentional states) should remain after the first has been treated: for surely if one has explained the causal basis of intentionality, how mental states reliably correlate with entities in the environment, one has already assumed that mental states are the kinds of entities which enter into causal relations. So what can the further question about their efficacy amount to?

Jacob thinks that the further question concerns the efficacy of the intentional *properties* of intentional states. He makes a distinction between two claims one can make about mental causation: the first is that intentional states are causes; the second is that intentional states are causes in virtue of their intentional properties (pp.32-33). He then claims that while the naturalistic theory of content explains the efficacy of mental states,

it leaves open whether they are efficacious in virtue of their contents. The role of the second half of the book, then, is to settle this question.

I would have liked a little more detail on the metaphysics of this conception of the problem of mental causation. To make the distinction between the two claims about mental causation, we need a distinction between a state and its intentional properties. This suggests either that states are particulars, or that intentional properties are higher-order properties of mental states—that is, properties of mental properties. It's not clear which of these options Jacob wants to adopt, and it is not clear whether in fact the distinction ought to be made for mental states as a whole, given Jacob's view of mental states.

On the first conception, states are entities rather like Davidson's events: unrepeatable particulars, the semantic values of certain singular terms, and the bearers of (first-order) properties. The trouble with this, from Jacob's point of view, is that although there are very good reasons for believing in events construed as particulars, there are few good reasons for thinking that all mental phenomena are events in this sense. A belief, for instance, is implausibly construed as something which *happens* to someone. Belief, on most views, is a disposition, and a disposition is not an event, though its manifestations are. So acquiring a belief may be an event, but the belief thus acquired isn't.

However, suppose we restrict ourselves to those mental phenomena which are uncontroversially events. It certainly is possible then to raise the question of whether this property or that property of an event is efficacious in bringing about a certain effect. In fact, this is the form in which the question of mental causation is sometimes raised as a problem for Davidson's view of the mind.<sup>1</sup> Jacob discusses this in a number of places (pp. 169-173; 212-215) but nowhere indicates that the problem can be raised in this form only for events construed as particulars (and not for all mental phenomena). This suggests

---

<sup>1</sup>See for example, the discussions in part one of J. Heil and A. Mele (eds.) *Mental Causation* (Oxford: Clarendon Press 1992).

that Jacob is really thinking of the second conception of the metaphysics of mental states: states are instances of properties, which themselves have properties.<sup>2</sup>

But this second conception is hard to square with other views Jacob holds. This much is clear: the view is committed to states, construed as instances of properties. Mental states are instances of mental properties. (States are therefore what some call ‘facts’, what others call ‘states of affairs’, and what Kim calls ‘events’.) So my belief that pigs fly is an instance of a property: it is natural to say this is the property of *being the belief that pigs fly*. If my belief is an instance of this property, then the question about the efficacy of the belief is precisely the question of the efficacy of *this* property (the one picked out by the italicised phrase above). But this can’t be how Jacob sees it, since he thinks that the question of the efficacy of the belief state is distinguishable from the question of the efficacy of its properties (specifically, its semantic/intentional properties). This is why he says that the content of a belief is a property of the belief (p.35). This is to reject both the traditional view that a belief is a *relation* to its propositional content—the view that gives talk of ‘propositional attitudes’ its sense—and also the alternative view that the belief that pigs fly is a determinate of the determinable *belief* (since determinates are not properties of their determinables, nor vice versa).

In addition, it is now hard to see what Jacob’s ‘token physicalism’ (pp.33ff) amounts to on this picture. For according to the standard view of the identity conditions of states, states S1 and S2 are ‘token identical’ when they contain the same constituent things, properties and times. So if mental states are instantiations of properties, then the identity conditions of the mental states are the identity conditions of their constituent properties. On this view, then, the ‘token’ identity theory collapses into the ‘type’ identity theory. But on a type identity theory there is, arguably, no problem of mental causation at all.<sup>3</sup>

---

<sup>2</sup>This seems to be confirmed by his endorsement of the view that mental properties are functional properties (p.216), in the sense of being properties which consist in the having of certain other properties. Jacob also says that he assumes that the relata of the causal relation are events (p.213) but since he is neutral on what events are—i.e. on whether they are particulars in Davidson’s sense—this commitment alone does not settle the issue.

<sup>3</sup>See Tim Crane, ‘The mental causation debate’ *Proceedings of the Aristotelian Society* 1995.

Setting these metaphysical worries aside, let us turn to Jacob's conception of the problems of mental causation. He plausibly argues that the worry about the inefficacy of the intentional comes from two sources (p.222): the first is that having a particular intentional content is not a 'basic' property of a mental state (i.e. it is not a physical property); the second is that since externalism is true, having content is not a local property of a state. The first problem of mental causation is that mental properties might be 'preempted' by more basic causes; the second problem is that only local properties are causes, and intentional properties are not local, given the truth of externalism.

Jacob proposes solving the first problem by distinguishing between the kind of causal explanations which are given in the case of intentional states and in the case of more 'basic' states. Suppose I have a headache and I take an analgesic (eg aspirin) to suppress the pain. So the suppression of my pain may be explained by the fact that I ingested an analgesic. But the aspirin relieves my pain because it contains acetylsalicylic acid. The presence of acetylsalicylic acid is a more basic property than the functional property of being an analgesic, so given the causal priority of basic properties, it seems that the efficacy of the property of being an analgesic is 'screened off' (or in Jacob's somewhat unhelpful terminology, 'pre-empted'<sup>4</sup>) by the more basic chemical property. Intentional properties are construed, like the property of being an analgesic, as functional properties, so the analogous question about their efficacy can be raised: intentional properties cannot be efficacious because they are screened off by their non-intentional realising properties.

Following Jackson and Pettit, Jacob distinguishes between *process* and *programme* explanations. In explaining why the acid relieves the headache, we pick out the acid as the causally efficacious element in a causal process. But this does not render the explanation in terms of the analgesic redundant: for we can say that invoking the

---

<sup>4</sup>Unhelpful because cases of causal preemption are normally construed diachronically, with the preempted cause being an event or state which would have caused an effect were it not for the fact that something else got there first. (See Peter Menzies, 'Probabilistic causation and the preemption problem' *Mind* 1996.) It is not normally assumed in setting up preemption problems that there is something about the nature of preempted causes which prevent them from being genuine causes. In Jacob's cases, however, the 'preempted' properties are simultaneous with those which do the preempting; and it is in the nature of these functional properties, as he construes them, that their efficacy is problematic.

property of being analgesic ‘programmes’ another more basic property to bring about the effect. This is the form of explanation which we offer when, for example, we explain why a cup is broken by saying that *someone* dropped it, even though there is a more specific explanation available—say, that Vladimir dropped it (p.218).

The trouble with this is that it does not offer a theory of causation or causal explanation, but merely an image: the image of one property ‘programming for’ the presence of another. That is, we explain how x’s having property F makes it the case that y is G in terms of the fact that x’s having property F requires x to have some other property, F\*, and F\* causes G. Put this way, who can deny that such causal transactions go on? This is what Jackson and Pettit mean by ‘programming’. But why should the relation between F and G not be a causal relation? The main answer Jackson and Pettit give is that genuine causal efficacy is reserved for more basic (or even for the most basic, fundamental) levels of existence: the microphysical level.<sup>5</sup> Yet this is not a thesis about causation, but about the nature of the physical world. As far as causation is concerned, we have little difficulty in seeing how F can cause G. If we put the physicalist thesis about the causal nature of basic properties to one side, the relation between F and G can plausibly be seen to satisfy the uncontroversial connotations of causation: F makes G more probable than it would otherwise have been, F is necessary or sufficient in the circumstances, a lawlike relation is entailed between F and G, and so on.

But in any case, it is not clear how Jacob’s approach here really addresses the issue. For the problem about ‘screening off’ is not a problem about the causal explanations we give of phenomena, but a problem about causation. The problem was that we have two competing causes for the same effect, and given the causal priority of basic properties, it looks like the intentional properties do not make anything happen. To solve this problem, we need to know *either* whether we have overlooked something about mental properties, *or* whether one of our assumptions (e.g. about the efficacy of basic

---

<sup>5</sup>Jackson gives another answer: the relation between dispositions/functional properties and their effects is conceptual, and therefore cannot be causal. See Jackson, ‘Mental properties, essentialism and causation’ *Proceedings of the Aristotelian Society* 1995. For scepticism about this attitude to properties and causation, see Tim Crane, ‘The efficacy of content: a functionalist theory’ in Jan Bransen and Stefan Cuypers (eds.) *Human Action and Causality* (Dordrecht: Kluwer, forthcoming).

properties) is incorrect. It cannot help us to be told that we can give equally useful and informative *explanations* of the same phenomenon in terms of intentional and in terms of non-intentional properties: for, in broad outline, we knew this already. The original worry was not about how many kinds of explanation we can give. To assume this is to set the problem of mental causation up on much stronger premises than are usually thought necessary: that is, to assume that there cannot be more than one explanation of the same phenomenon. Although some people (notably Kim<sup>6</sup>) have assumed this, they are best interpreted as using the term ‘causal explanation’ as a variant for ‘causation’.<sup>7</sup> But Jacob explicitly distinguishes between causation and causal explanation (p.214), and explicitly says (p.221) that he is understanding ‘explanation’ in an epistemic sense. So this way out is not open to him.

The trouble with Jacob’s solution to the screening off problem, then, is that by appealing to the different kinds of explanation we give of phenomena, and giving the notion of explanation an epistemic interpretation, Jacob fails to engage with the metaphysical worries which generate the problem of mental causation. (And if, on the other hand, the distinction between process and program explanation is taken as a distinction between kinds of causation, then my complaint would be that we have not been given a theory, only an image. But in any case, as I observed above, Jacob does not take this route.)

The central question is: how do the causal explanations we give pick up on the causal features of reality? Jacob does say that his two explanations (process and programme) do track the same causal process:

different causal explanations, tracking one and the same causal process, may deliver different pieces of information about the causal process which they track.

By referring to the causally efficacious chemical property of the pill, the chemical explanation reveals the actual process of pain suppression. By mentioning the pill’s functional property of being analgesic, the functional explanation supplies information about a class of possible unrealized chemical processes. (p.221)

---

<sup>6</sup>See Kim’s idea of ‘explanatory exclusion’ in ‘Mechanism, purpose and explanatory exclusion’ in *Supervenience and Mind* (Cambridge: Cambridge University Press 1993).

<sup>7</sup>For this way of talking about causal explanation, see David Owens, *Causes and Coincidences* (Cambridge: Cambridge University Press 1989).

Two things about this passage deserve comment: first, the picture suggests that there is one causal process which may be described ('tracked') in two ways. This, plus Jacob's distinction between causation and causal explanation, is reminiscent of Davidson's metaphysics, where the same relations between the same particular events may be given different explanations in incommensurable vocabularies. And Jacob sometimes suggests that he endorses some of Davidson's views (e.g., pp. 161-162). But on Davidson's picture, contrary to what many of his critics believe, there is no problem of mental causation, so the distinction between program and process explanation is not required to solve it (though it may be a useful distinction between kinds of explanation—this is a separate issue).<sup>8</sup>

Second, the programme explanation is described as giving information about a 'class of possible unrealized processes'. Maybe so; this seems like a good way to describe part of what is going on. But what then is the rationale for calling this explanation a causal explanation? Whatever else is true in the controversial world of theories of causation, this much seems correct: 'A caused B' cannot be true unless A and B actually exist. So citing classes of unrealized possibilities, however informative, cannot be citing causes. I am sceptical then, about whether all the metaphysical structures picked out by Jackson and Pettit's programme model are really cases of causation.

When it comes to the second problem of intentional causation—the problem raised by externalism—Jacob adopts a different strategy. He argues (against Jackson and Pettit, but correctly, in my view) that the programme/process distinction cannot help us in explaining how broad mental states are efficacious. Instead, to solve this problem, he endorses two views defended in Dretske's 1988 book, *Explaining Behavior*. The first view is that a piece of behaviour should not be identified with the bodily movement which is the causal product of certain brain states; rather behaviour is the *process* of the brain states causing the bodily movement. The second view is that although broad

---

<sup>8</sup>For Davidson's response, see 'Thinking causes' in Heil and Mele (eds.) *Mental Causation*; for a defence of Davidson, see Crane, 'The mental causation debate'.



intentional states do not cause bodily movements in the way that internal brain states do (in the ‘triggering’ sense of cause) they do causally *structure* the behaviour conceived of as a process. Dretske’s solution thus involves distinguishing the explananda of psychological explanation—process, not product—and distinguishing the kind of causation which psychological explanation picks out—structuring, not triggering.

It is not hard to see how Dretske’s re-orientation of the traditional way in which this debate is set up would show how broad intentional states are causally efficacious (in his new sense). The question rather concerns whether the re-orientation is justified, and whether Dretske avoids solving the problem by convenient stipulations. Jacob does a good job of defending Dretske’s views here, but his defence raises a further question about his strategy. If Dretske is right, and his new conception of psychological explanation solves Jacob’s second problem of mental causation, then it also solves the first problem: for if intentional states cause behaviour in Dretske’s sense, then there is no question of them being screened off by more basic physical properties. The question simply doesn’t arise, since intentional states and the basic properties are not competing for the same explananda. So why does Jacob bother to endorse the Jackson/Pettit view, if Dretske’s more radical solution would solve both problems at one stroke? The problem here, like the other problems I have mentioned, may derive from Jacob’s (otherwise admirable) ecumenical approach to the philosophy of mind. *What Minds Can Do* develops and expands upon the views of its main influences—Fodor, Dretske and Davidson—but it seems to me that no coherent philosophy of mind is big enough to make all these philosophers’ views mesh in a satisfactory way.

Tim Crane  
Department of Philosophy  
University College London  
Gower Street  
London WC1E 6BT