

## Article

# Folk Belief and Commonplace Belief

FRANK JACKSON AND PHILIP PETTIT

Our position on eliminativism is very different from Barbara Hannan's as far as premisses go, though we agree with her overall conclusion that eliminativism is to be rejected. We hold: (i) that beliefs and desires (and the propositional attitudes in general, but we will follow the common practice of focussing on beliefs and desires), are, according to the folk conception of them, posits of an explanatory theory of behaviour, the theory known as 'folk psychology'; (ii) that, in consequence, neuro-science might show that there are no beliefs and desires; but (iii) that we know enough about the genesis of behaviour to know that this will not happen. The case for such a position calls for a book, not a discussion, but we will aim to say enough to indicate its possibility and appeal.

### 1. Beliefs and Desires are Explanatory Posits according to the Folk

The folk attribute beliefs and desires to human beings; they do not attribute beliefs and desires to stones; and there is no mystery about the reason for the difference. It lies in the fact that human beings behave in highly complex ways, whereas stones do not. There are, accordingly, it seems to us, two live options for what, in broadest outline, the folk conception of belief and desire might be. They might be patterns in the actual and possible behaviour-in-circumstances that creatures like us manifest. When I say that Smith believes that snow is white, I might be ascribing a certain character to his behaviour, actual and possible (not just actual, of course, because the folk are prepared to revise their belief

and desire ascriptions when they learn about what someone would have done, even when they know that the behaviour did not actually take place). Alternatively, the folk conception might be that beliefs and desires are underlying internal causes of the kinds of behavioural patterns that ground ascriptions of belief and desire. Either hypothesis makes sense of the famous role of belief and desire ascriptions in predicting behaviour-in-circumstances: prediction is facilitated both by the identification of projectible patterns, and by the identification of underlying causes. The prediction of the behaviour of an electron in a force field goes just as well when carried out by a phenomenalist about force fields, as it does when carried out by a realist about them.

How then do we choose between the two conceptions? Well, if our topic is the folk conception of belief and desire, a good procedure is to ask the folk. This is what we do when we ask them for their reaction to thought experiments; and their reaction provides strong evidence that their conception of belief and desire belongs to the 'underlying causes' category. Consider, for instance, Christopher Peacocke's Martian marionette example.<sup>1</sup> We discover that certain human-shaped figures, indistinguishable from ourselves in terms of behaviour-in-circumstances, are being operated from Mars: their heads are receiving stations for instructions from Mars as to how to behave. The universal response to this example is to say that the marionettes do not have beliefs and desires. But then belief-desire talk is not mere labelling-behavioural-patterns talk. We are, rather, taking on a substantial commitment to the nature of the underlying causes of behaviour—and, in particular, as revealed by this example, we are committing ourselves to the underlying causes not being of the nature of a mere receiving station. A more complex case that points in the same direction is Blockhead, the 'creature' which behaves in just the way a typical person does in all actual and possible situations it might encounter in the course of a life, but does so, not by processing information in anything like the way we do, but by following an *enormous* implanted look-up tree. Blockhead 'gives' the behavioural answers to life's actual and possible challenge's in the same way as someone who cannot work out square roots answers questions about square roots, namely, by recourse to a table. Its behaviour, unlike that of the marionette, is internally generated to the same extent as ours, but is generated in such a way that folk intuition delivers that it does not have beliefs and desires.<sup>2</sup>

Also, there are the reactions of the folk to cases involving subjects who are completely paralysed. The folk allow that such subjects can, and sometimes do, have beliefs and desires, and yet they do not display the

<sup>1</sup> See the final chapter of Peacocke (1983).

<sup>2</sup> See Block, 1981, for a persuasive account of the example. The focus of the argument in his paper is on whether Blockhead (we do not know who first coined this name for the creature Block describes) is intelligent—it isn't—but it is equally clear that it has no mental life at all.

behavioural patterns that ground our ascriptions of belief and desire. The obvious explanation is that what matters, according to the folk conception, is that such subjects have inside them the underlying states which would, in normal circumstances, cause the relevant behaviour.

It thus seems to us that there is a strong case for the view that the folk conception of belief and desire is that of internal states playing *inter alia* certain kinds of causal-explanatory roles with respect to behaviour. But, equally, there is a strong case for the view that neuroscience, when supplemented with the appropriate information about a subject's surroundings, can in principle, and maybe one day will in practice, explain each and every aspect of behaviour-in-circumstances. How then can we resist the conclusion that it is at least an open possibility that neuroscience's explanation will displace the predictively impressive, but notoriously fallible and restricted in scope, explanations of behaviour in terms of belief and desire?

Some respond to this argument by urging that neuroscience cannot explain everything about behaviour-in-circumstances. They urge that there are facts about behaviour-in-circumstances that the positing of beliefs and desires can explain, and which cannot, not even in principle, be explained in terms of neuroscience. We disagree, but will not enter that familiar debate here. We want to describe (in outline) how to resist the conclusion at the same time as granting the premise that neuroscience can do it all in principle.

We start with some remarks on the example of imprinting.

## 2. Learning from (Filial) Imprinting

Chickens are disposed to keep company with the first thing of a suitable sort that they see after hatching. Usually it is the mother hen that they see first, but sometimes it is a small dog, the experimenter, or whatever. They are said to imprint on the first thing that they see, and their behaviour is explained in terms of their having imprinted on the mother hen, the experimenter, or whatever.

Although we ascribe imprinting on the basis of observation of behaviour-in-circumstances, what we ascribe, or are in a position to ascribe, goes well beyond the facts about behavioural patterns. We know, for instance, that: the chicken's initial sighting lays down a persisting trace inside the chicken, otherwise its following behaviour would fade away quickly; that the nature of the internal trace that is laid down is a function of the nature of the thing first seen, otherwise it would not be able to discriminate between the thing first seen and things seen subsequently; and that the trace laid down is causally connected to the ways the legs and the head of the chicken operate, otherwise the information being carried by the trace inside the chicken would be irrelevant to the observed operation of its head and legs in sustaining the accompanying behaviour.

What is important for our purposes about this example is that it would clearly be a serious mistake to argue as follows about imprinting: 'The imprinting story is a little theory about the causation of the chicken's behaviour. The predictions it yields are much better than guessing, but much worse than those that would be yielded by a comprehensive neuroscientific story about the chicken's internal workings combined with an account of the relations of these internal workings to its environment. In particular, imprinting makes some wrong predictions, and, most strikingly, is silent about a whole range of behaviours that chickens manifest; and, worse, there seems no way of naturally building on the "theory" of imprinting to handle these facts about the behaviour of chickens. The theory of imprinting shows all the signs of a degenerating research program.'<sup>3</sup>

The situation with the theory of imprinting is, rather, that, although in principle the theory of imprinting could be overthrown by the discoveries of neuroscience, we know perfectly well that it won't be. The behavioural evidence is too strong for that to be more than abstractly possible. As we might put it, the commitments to how things are inside the chicken are commitments to commonplaces. Neuroscience will not displace these commonplaces. Rather, it will bear them out: it will reveal how the traces are stored, how they manage to be informationally sensitive to diverse initial sightings, and how they link to the muscles that control the head and legs of the chicken.

Our view is that, as it is for theory of imprinting, so it is for the theory of belief and desire. The theory of belief and desire does involve substantial commitments about how things are inside subjects, but the commitments are like those that feature in the theory of imprinting—ones that are *very* plausible given what we know concerning how subjects behave in circumstances. A proper defence of this position is beyond the scope of this discussion. What we will do is describe how we think the defence should go in outline. We will also allow ourselves some remarks on how the defence should *not* go.

## 3. Commonplace Psychology

The question, Are there Ks? can be divided into two questions: What does it take for there to be Ks? and, Is what it takes in fact the case? There are electrons because what it takes for there to be electrons is in fact the case. There are no unicorns because what it takes for there to be unicorns is not the case. What we are seeking, then, is an account of what it takes for there to be belief and desire according to which it is very plausible

<sup>3</sup> It would also be a mistake to think that the broad nature of imprinting meant that it was unsuited for the explanation of behaviour, but that is another story.

that what it takes is in fact the case. How might we obtain such an account?

We might invent one, especially tailored to do the job. But then what would be the interest of the result? We would merely have shown that, according to our cooked-up account of belief and desire, there are beliefs and desires. What is required, what is at issue (as opposed to what we, or others, might like to say was at issue in order to make our, or their, case), is whether there are beliefs and desires according to something like the folk concept of belief and desire. Now, it seems to us, that the folk concept of belief and desire is captured by commonsense functionalism, for commonsense functionalism is constructed by putting together all the things the folk hold to be most certain about belief and desire. But there could, of course, be no prior guarantee that there are beliefs and desires according to the folk concept so elucidated. Some of what the folk hold to be most certain about belief and desire might not be true of anything. What we can do, however, is to construct concepts of belief and desire which are at least close relatives of the folk concepts, and for which there is a prior guarantee that there are beliefs and desires according to the concepts so elucidated. The idea is to appeal to a close relative of folk psychology, which we will call *commonplace psychology*.

Commonplace psychology is folk psychology minus what might be seriously doubted. Following David Lewis, we can think of folk psychology as the long conjunction of everything we folk take for granted, explicitly or implicitly, as pretty much common opinion about mental states.<sup>4</sup> On this picture we think of folk belief as the state which occupies, or near enough occupies, the 'belief' place in this long conjunction. In outline, the story runs as follows. Let **M** be the long conjunction. Replace each distinct mental property term by a distinct variable to give ' $M(x_1, \dots, x_n)$ '. Then ' $(Ex_1)(Ex_2) \dots M(x_1, x_2, \dots)$ ' is the Ramsey sentence of **M**. It gives the content of folk psychology.<sup>5</sup> And folk belief is simply the state which plays the belief role in **M** in the sense that if ' $x_b$ ' replaced 'belief' in **M**, then I am a believer according to the folk concept if and only if  $(Ex_1)(Ex_2) \dots$  [I have  $x_b$  &  $M(x_1, x_2, \dots)$ ]. The content of *commonplace psychology* is given by the *cautious* Ramsey sentence of **M**. The *cautious* Ramsey sentence of **M** is simply ' $(Ex_1)(Ex_2) \dots M(x_1, x_2, \dots)$ ' with anything open to serious doubt deleted or modified to make it pretty much a truism: we might write it ' $(Ex_1)(Ex_2) \dots m(x_1, x_2, \dots)$ '. Then commonplace belief is simply the state which plays the commonplace belief role in **M**—that is, if ' $x_b$ ' replaced 'belief' in **M**, then I am a believer according to the commonplace concept if and only if  $(Ex_1)(Ex_2) \dots$  [I have  $x_b$  &  $m(x_1, x_2, \dots)$ ]. In short, commonplace belief is what occupies those of the roles

associated with belief in the folk theory which we cannot seriously doubt are occupied. In the same general way we could elucidate commonplace desire, and, indeed, commonplace hope, commonplace perception, and so on.

The issue about eliminativism can now be put as follows. There can be no doubt that there is commonplace belief and commonplace desire. That follows from the way we specified commonplace psychology in terms of the cautious Ramsey sentence. And their existence will count as there being belief and desire provided commonplace psychology is enough like folk psychology, provided, that is, we do not have to delete or modify too much of importance in the Ramsey sentence of **M** to get the cautious Ramsey sentence of **M**; for if we have to delete too much, we will not be giving a negative answer to eliminativism, but rather will be changing the subject. Our conviction—unargued here—is that commonplace psychology is quite similar to folk psychology, and, in particular, similar enough for commonplace belief and commonplace desire to count as belief and desire.<sup>6</sup>

#### 4. An Easier Way?

Our outline of a defence of the existence of belief and desire assumed that the folk concept of belief and desire is that of states occupying the roles specified in folk psychology. That was why the question of the similarity of commonplace psychology to folk psychology loomed so large. But William Lycan has suggested that (Lycan, 1988, p. 32):

As in Putnam's examples of 'water', 'tiger', and so on ... the ordinary word 'belief' (qua theoretical term of folk psychology) points dimly towards a natural kind that we have not fully grasped and that only mature psychology will reveal. I expect that 'belief' will turn out to refer to some kind of information-bearing inner state of a sentient creature, ... but the state it refers to may have only a few of the properties usually attributed to beliefs by commonsense.<sup>7</sup>

If Lycan is right, if the folk concept of belief is that of a natural kind in common between certain exemplars—us and those like us, presumably—which we have tagged or baptised with the term 'belief', rather than being that of a state satisfying a good many of the descriptions we associate with belief, then the existence of folk belief will not depend on our folk

<sup>4</sup> See, e.g. Lewis, 1972.

<sup>5</sup> Or, if uniqueness of realisation is part of the folk story, ' $(Ex_1)(Ex_2) \dots (y_1)(y_2) \dots [M(y_1, y_2, \dots) \text{ iff } x_1 = y_1 \text{ \& } x_2 = y_2, \dots]$ ' gives the content of folk psychology, but we will neglect this complication here.

<sup>6</sup> For some arguments, see Jackson and Pettit, 1990a, 1990b. In (1990a) we argue that the folk assumption sometimes called 'propositional modularity' does not involve non-commonplace assumptions about our insides.

<sup>7</sup> Putnam's account is set out in, e.g. Putnam, 1975. See also Kripke, 1980.

theory being largely correct, and the controversial question of the similarity of commonplace psychology to folk psychology could be set aside. Things would be that much easier for the defender of belief (and desire).

We think, however, that there are serious problems for a tag view of our concept of belief (and desire, but let's focus on belief). Exemplars of water have a leading candidate—H<sub>2</sub>O—to be the natural kind they all belong to. Although there are many similarities between the exemplars of water, only one corresponds to a natural kind. It is in virtue of this fact that the term 'water' picks out that natural kind on the Putnam–Kripke story, and it is in virtue of this fact that the theory has its attractions. But what natural kind membership do exemplars of believers like you and me share which might plausibly be held to be that which the term 'belief' tags?

We share membership of the biological natural kind: living human being. But if the term 'believer' tags that natural kind, anything which is not a human being cannot be a believer. (Tagging is rigid.) And one of the few things almost everyone agrees upon in these stormy waters is that we should not be chauvinists. Dogs could have beliefs. Similarly, it would be a mistake to hold that 'believer' tags the kind: warm-blooded, carbon-based creature. Silicon-based, cold-blooded Vesuvians might have had beliefs.

Lycan talks of 'some kind of information-bearing inner state of a sentient being', and it is plausible that there are significant, that is, non-gruesome, similarities between the information-bearing states inside us at a level higher than their neurophysiological nature. But, of course, there is an awful lot of information processing going on inside us, very little of it having anything essentially to do with belief. The information processing that goes on in our stomachs is no doubt rather similar to that which goes on in the stomach of a dog, but that is irrelevant to whether or not the dog and we are alike in being believers. The relevant information processing must be that underlying the distinctive behavioural capacities common to the beings we suppose to be believers. Let us suppose, then, that there is such a distinctive style of internal information processing, that it marks off a natural kind on some account of that elastic notion, and that the term 'belief' tags it. Something along these general lines seems to be the most plausible development of a causal-historical account of the semantics of 'belief'.

Nevertheless, there is trouble. For there are more forms of chauvinism than the neurophysiological, chemical and biological versions so extensively criticised in the literature. Why should creatures whose style of internal information processing is markedly different from ours—that is, from the exemplars of believers—thereby be debarred from the club? Plausibly, there are styles of internal information processing that are inconsistent with being a believer—the style that goes on inside Blockhead would be an example—but it would be unduly chauvinistic to insist that in order to count as a believer a creature must solve the information

processing problems a world sets its inhabitants in the same general way that we do, or indeed in anything much like the way we do. Monocular silicon-based Vesuvians would no doubt process information about the location of the objects around them carried by the light impinging on them very differently from the way we binocular carbon-based Earthians do, but any science fiction buff knows that our concept of belief does not thereby preclude the Vesuvians from having beliefs about, say, the location of the objects around them.

*Research School of Social Sciences  
Australian National University  
Canberra, ACT 0200  
Australia*

### References

- Block, N. 1981: Psychologism and Behaviorism. *Philosophical Review*, 90, 5–43.  
 Jackson, F. and Pettit, P. 1990a: Causation in the Philosophy of Mind. *Philosophy and Phenomenological Research*, 50, Supplement, 195–214.  
 Jackson, F. and Pettit, P. 1990b: In Defence of Folk Psychology. *Philosophical Studies*, 59, 31–54.  
 Kripke, S. 1980: *Naming and Necessity*. Oxford: Basil Blackwell.  
 Lewis, D. 1972: Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy* 50, 291–315.  
 Lycan, W.G. 1988: *Judgement and Justification*. Cambridge: Cambridge University Press.  
 Peacocke, C. 1983: *Sense and Content*. Oxford: Oxford University Press.  
 Putnam, H. 1975: The Meaning of 'Meaning'. In *Mind, Language and Reality*. Cambridge: Cambridge University Press, 215–71.