

# Is Folk Psychology Different?

JONATHAN KNOWLES

PHILOSOPHY DEPARTMENT (MODULE 1-CENTRE)

NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY, 7491 TRONDHEIM, NORWAY

E-MAIL: JONATHAN.KNOWLES@HF.NTNU.NO

**ABSTRACT:** In this paper, I seek to refute arguments for the idea that folk psychological explanation, i.e. the explanation of actions, beliefs and desires in terms of one another, should be understood as being of a different character than ordinary scientific explanations, a view defended most prominently in analytical philosophy by Donald Davidson and John McDowell. My strategy involves arguing both against the extant arguments for the idea that FP must be construed as giving such explanations, and also against the very notion of such a different kind of explanation. I argue first that the in some sense a priori and conceptual nature of folk psychological principles does not support the idea that these are other than empirical generalisations, by appeal to recent nativist ideas in cognitive science and to Lewis's conception of the meaning of theoretical terms. Second, I argue that there is no coherent sense in which folk psychological explanations can be seen as normative. Thirdly, I examine the putatively holistic character of the mental and conclude that that too fails to provide any cogent reasons for viewing folk psychological explanations as different from other kinds of explanation.

**KEYWORDS:** Folk psychology, explanation, conceptually true, normative explanation, holism, idiographic explanation, simulation theory.

## *Introduction*

A philosophical issue at the heart of many debates in philosophy of mind and scientific methodology in the 20th century has been the status of psychology as a science. Can we study the human mind and its manifestations using essentially the same methods and concepts as employed in natural science – deriving hypotheses couched in causal and nomic terms

## *Is Folk Psychology Different?*

through observation, theoretical elaboration and experimentation? Or should psychology instead be viewed as a branch of the humanities and/or of social studies – even if as such it amounts to nothing more than the insights we gain from poets, authors and dramatists (and perhaps those who study them)? For those with sympathies for the recent developments in cognitive science – in linguistics, vision research and developmental cognitive psychology *inter alia* – this philosophical issue will be regarded (along with so many others in the past) as superseded by the fact that scientific research into, in this case, the structure and origin of our mental capacities has met, and continues to meet with success. And indeed there would seem to be no doubt that a science worthy of the name ‘psychology’ is today pursued in the manner of a progressive research programme, alongside biology, chemistry, physics and the many special sciences that concern themselves with particular aspects of the physical world. Nevertheless, a question remains as to whether there is not some dimension to, or aspect of our mind which still evades the theoretical grasp of cognitive science – and not merely for uninteresting, ‘contingent’ reasons, such as the immaturity of the discipline.

That this might be the case is something which even two arch-proponents of cognitive science, Jerry Fodor and Noam Chomsky, have often averred (cf. e.g. Fodor 1983, 1998: §§ III-IV; Chomsky 1980: 6-7, 1988: chapter 5). However, it is significant that Fodor and Chomsky do not see the limitations of cognitive science as indicating the availability of an alternative understanding of the mind. Nor do they suggest that this understanding would have any special characteristics were it, perhaps *per impossibile*, to be available to us. Theirs is the largely negative point to the effect that we may – indeed, that it is likely on general theoretical grounds that we will – always remain in the dark with respect to certain central mysteries about the mind, such as the problems of consciousness, free will and scientific belief formation.

This piece will not seek to pronounce on the possibilities that Fodor and Chomsky aver. Nor will it seek to argue that a study of the mind definitively is, or should be, through and through scientific. Instead it will be concerned with those who hold, not just that certain aspects of our mind demand an understanding other than that which cognitive science can

offer, but that this understanding is (a) one which involves giving substantive but distinctly *non-scientific* explanations of mental phenomena, and (b) something we are naturally in possession of. This idea – which in contemporary analytic philosophy stems from Davidson’s writings on the anomalism of the mental, and has found its most eloquent exponent in John McDowell – has been connected to the debate about what has come to be termed ‘folk psychology’.<sup>1</sup> Folk psychology is often described as the scheme of explanation employed by all ordinary, mature human beings to explain and predict others’ behaviour by attributing them mental states, such as beliefs, desires, pains, emotions, and so on. More narrowly, but also more directly relevantly for our concerns, folk psychology (henceforth FP, and understood for the rest of this essay in the following sense) involves, in particular, attributions of propositional attitudes to explain and predict belief- and desire-formation and intentional action. Davidson’s and McDowell’s – and many others’ – view of FP is that its explanations are of a different character than scientific explanations – which means, amongst other things, that they are not deductive-nomological; they are instead in some sense *normative*. The view of most contemporary cognitive scientists and philosophers sympathetic to cognitive science is, by contrast, that these explanations are essentially parallel to those found in other sciences, or are proto-versions of such explanations. Further, for these latter thinkers, though giving FP-explanations may be practically indispensable for human beings, and thus in some sense rational whatever we get to know about the mind scientifically, the ultimate fate of FP as a true, explanatory scheme depends on developments in cognitive science, both conceptual and empirical (i.e. on whether we can wring FP-explanations into some usefully precise predictive and explanatory model, and on whether this turns out to be substantially correct, fruitful etc.). For those who hold FP to have an essentially normative character, developments in science

---

<sup>1</sup> Or – less commonly though perhaps also less tendentiously – ‘common sense psychology’. As several authors have pointed out, the epitaph ‘folk’ sparks associations with folk sciences of other kinds, such as folk physics. This is regarded as unfortunate insofar as folk physics is generally regarded as both (i) involving a form of scientific understanding (albeit primitive) (ii) false – neither of which is uncontroversially the case with folk psychology. Nevertheless, since this point has been made so often and the epitaph ‘folk psychology’ retained in philosophical discussions of the nature and tenability of folk psychology, this mild inappropriateness would seem to be assuaged. I will in any case use ‘folk psychology’ in the following without implying either (i) or (ii) holds of folk psychology.

will have no impact on the truth content of our folk explanations of others' behaviour; they are in this sense *autonomous* from science. Yet this is not because FP concerns a sphere of the mental that is essentially ineffable, or which simply evades scientific understanding.

I will be arguing against the view that FP gives explanations of an essentially different character than those found in science in the way assumed by Davidson, McDowell and those influenced by these two. My strategy will involve arguing both against the extant arguments for the idea that FP *must* be construed as giving such explanations, and also against the very notion of such a different kind of explanation. My conclusion will be, not that FP is scientific – for it may be that it, either essentially or contingently, falls short of what we expect of a scientifically respectable theory – but that there is no cogent case for thinking that, when it comes to explanations, FP is in any other business than giving, at least in some loose sense, scientific ones.

I will not be making many specific assumptions about the criteria for scientific explanation or explanation in general. For simplicity's sake I will be assuming that there is no principled distinction between scientific and non-scientific explanation generally – that is, with the possible exception of FP-explanation. This is not a substantive commitment: if, for example, explanations that do not essentially involve laws are deemed non-scientific, then I would concede that FP-explanations perhaps are not scientific, even though they are not in any way special, as Davidson and McDowell hold. To put it another way: it is no part of my remit to show that FP-explanation is more than proto-science, or that it is scientific in a way many ordinary, everyday explanations are not. Connectedly, nothing of what I want to maintain presupposes that a classical deductive-nomological – or a more nuanced but still essentially monolithic – understanding of explanation must be correct. Having said all that, the views I discuss tend to contrast the insight offered by FP-explanation with that given by something like D-N-explanation, tacitly assuming that the latter is definitive of scientific explanation. I will argue that these views present no coherent alternative to a style of explanation which invokes empirical generalisations. I will also argue that rejecting a

‘subsumptive’ model of explanation fails to allow one to distinguish folk psychological explanation from other kinds of non-law-involving explanations.<sup>2</sup>

My conclusion – that there is no reason to think there is some aspect of the mind amenable to a special kind of non-scientific explanation – is, if correct, clearly of utmost significance. To begin with, it sharply delimits the implications of *hermeneutical* methodology for the humanities and social sciences: though there may be special methodologies for use in interpretation, these do not have implications for the explanation of human action, as many have supposed (this might in turn cast doubt on the significance of the former idea, though I will have nothing to say about interpretation *per se* in this piece – see the Conclusion.) Secondly, the question of the scientific status of psychology is significant because, insofar as systematic methods for investigation are consonant with a subject matter, the idea that humans and their minds demand special forms of explanation in relation to the rest of the natural world could be taken as indicating that, in spite of all the advances that have been made in science, there is still something special about us – in the sense, not merely that we are unique, for all forms of life are that, but rather that we are *uniquely* unique: that what makes us special is itself something special. In this way, the question about the methods of psychology reflects the vestiges of an old-age and more metaphysically charged debate concerning the place of man and mind in the natural world. If it turns out there is no special, non-scientific methodology for psychology, we should be prepared to give up the idea of our unique uniqueness.

The paper is divided into three main sections. The first considers Davidson’s views on FP-explanation, especially as these have been interpreted and defended more recently by Simon Evnine, and argues that this defence gives no reason to view FP as anything other than a (proto-)scientific theory. Section 2 considers McDowell’s views (which may be viewed as an elaboration of suggestions in Davidson), including his most recent writings on the idea of

---

<sup>2</sup> Note further that nothing of what I have to say should be understood as being sympathetic to the idea that FP should be seen as meshing with a computational-cum-representational theory of mind, as many take to be the case (cf. Fodor 1987: chapter 1; also Pickering & Chater 1995). I am in fact opposed to the idea that folk psychology, or any suitably revised intentional psychology, requires vindication by an implementational theory from cognitive science (cf. Knowles 2001), though nothing of what I have to say here depends on this.

*second nature*. I adduce considerations that militate against the coherence of McDowell's views. Section 3 takes up the problem of holism in relation to explanation in FP (and, briefly, cognitive science). I will argue that this aspect of the mind does not warrant thinking of FP-explanations as a substantive variety, at least divorced from the ideas already discussed. I conclude by clarifying my discussion in relation to various subsidiary ideas in the literature concerning the nature of FP-explanation.

*1. Davidson and the anomalism of psychology*

Davidson's views on physicalism and the nature of psychology are well-known, controversial and still widely discussed. He remains an important point of reference in the philosophy of psychology, though perhaps not now, at least chiefly, in virtue of his anomalous monism – a view which as the years have gone by seems to have taken the place of a bi-product of more interesting commitments. That which we are concerned with here is his view that there are no psychological laws in the way in which there are physical laws, where by psychological laws is meant, in the context of our present discussion, laws of folk psychology: generalisations relating beliefs, desires and actions to one another (henceforth 'FP-laws'; cf. Churchland 1981 for an opposing view and some examples of putative FP-laws). In the test of time, it is this idea, rather than the anti-reductionism that occupies central stage in 'Mental events' (in Davidson 1980) that has divided him most clearly from his opponents in the philosophy of mind. This interpretation is underscored, and Davidson's position generally clarified, in a more recent article (Davidson 1987).

For Davidson, FP-explanations are not deductive-nomological, they do not operate by 'subsuming' instances under empirical generalisations.<sup>3</sup> They are not scientific, and cannot

---

<sup>3</sup> In fact, Davidson has gone down in print as saying that FP-explanations can be viewed deductive-nomological (cf. his 'Hempel on explaining action' in his 1980), though he certainly would demur at the idea that, if we do this, it is a form of explanation by subsumption under empirical, psychological generalisation. In particular, he holds, first, that one can always redescribe the relevant events in physical terms such that a strict law covering cause and effect may, in principle, be uncovered; and, second, that even restricting oneself to psychological vocabulary, trivial deductions of facts about people acting in certain ways, e.g. that X eats a sausage at time t in situation S, may be given from a 'nomological' premise like 'If X is in situation S at time t and believes the

be rendered otiose by scientific advances – they are in this sense autonomous. Exactly how they do operate is a somewhat vexed question, and Davidson’s account involves rather vague talk of a ‘constitutive ideal of rationality’ as a guiding principle. Much of this and the following sections will discuss what this might involve. However, Davidson is perhaps most celebrated for his negative arguments in this area, in particular the argument that there are no psychological laws, that is, FP-laws. There are no strict, exceptionless generalisation of the kind we find in physics, or at least which physics aspires to, linking beliefs, desires and actions to one another. This is argued for by referring to particular cases, and also by appeal to the holism of the mental (cf. Davidson 1980 *passim*).

However, that there are no such strict laws is common ground between Davidson and many of his opponents on the nature of FP. Most prominent here is Jerry Fodor (cf. Fodor 1974, 1998: chapter 2), who has made much of the idea of *ceteris paribus* laws, and the fact that many special sciences use such laws – so why not FP? Davidson has now conceded that FP-explanations presuppose rough-and-ready generalisations of some kind, both in the light of Fodor’s observations and as a result of pressure (from, amongst others, Føllesdal 1985) to give an account of what makes the citation of a certain belief and a certain desire an informative, non-trivial explanation of a given action (cf. Davidson 1987, following page references to this article). However, he still denies that these should be viewed as laws in the way even those of other special sciences are (p. 45). Whilst there is no reason to think that the *ceteris paribus* clauses of the latter cannot in principle be spelled out – i.e. no reason to doubt that they may in principle be reduced to laws of fundamental physics – we know a priori that those of FP will never be able to be so reduced. Why not?<sup>4</sup> Because, says Davidson, FP-laws

---

sausage is on the table at S and t, she eats the sausage at S and t’ together with ‘X believes the sausage is on the table in S at t and is in S at t’. Since we are interested in FP-explanations, not physical ones, redescription in physical terms is irrelevant for us. As for the trivial deduction, Davidson’s point is precisely, in part, that the kind of ‘law’ mentioned does not represent a genuine empirical generalisation.

<sup>4</sup> This is of course a crucial question, yet it is striking that in discussions of the scientific nature of FP it rarely receives a straightforward answer beyond vague allusion to the ‘constitutive ideal of rationality’ (even the responses attributed to Davidson above have to be prised out of lots of round-about discussion). Thus William Child, in his defence of Davidsonian anomalism (Child 1994: ch. 2.1), concedes the parallel between special science and psychology in terms of use of *ceteris paribus* laws, but endorses Davidson’s line that there is a known, principled barrier to the removal of these in the case of psychology, ‘namely, the constitutive role of rationality and the uncodifiability of rationality’ (p. 62). The former idea is however not explained; the latter for Child entails the ideas i) ‘that there is no set of general principles from which, together with a specification of

represent *norms* that beliefs, desires and actions are identified in relation to: The principles are imposed by us and are constitutive of understanding what beliefs etc. are (cf. pp. 46-7).

This is a familiar formulation, even though its import is by no means wholly clear. The idea that FP-explanations are in some sense *normative* and *reflexive* (i.e. given relative to standards that are *ours*) seems close to the surface; this will be the topic of section 2.<sup>5</sup> In this section I will focus on what would seem to be a distinct interpretation suggested by Davidson's claim that a law such as "Someone who wants to crush a snail has a tendency to do what he believes will result in crushing the snail"... adds nothing to what we already understand if we know what a want or desire is' (p. 44). This suggests simply that we do not learn any more by knowing FP-laws than what we already know in understanding the concepts involved. They are known a priori through conceptual mastery, such that, were one to deny the generalisations, one could not be attributed the concepts they contain – rather as, were one to deny that all bachelors are unmarried, suspicion would arise as to whether one really possessed the concepts *bachelor* and *unmarried*. Hence, it might be claimed, the 'laws' of FP do not constitute an empirical theory as those of other special sciences do.

This proposal is somewhat underdeveloped in Davidson's paper, but it is a driving thought behind Simon Evnine's lucid exposition of Davidson's views on psychological explanation (cf. Evnine 1991: chapter 1). Moreover it is not, on the face of it, totally implausible. On the other hand, talking of FP-laws as conceptually true cries out for further elaboration, and other ideas we will say more about later – normativity and reflexivity – could

---

any agent's physical properties, we can derive a complete and detailed specification of her mental properties' and ii) 'that there can be no system of strict laws on the basis of which actions and other mental phenomena could be exactly predicted and explained' (p. 60). But the first of these ideas just records a commitment to anti-reductionism, which is common ground, whilst the second seems very much to simply assert that which one sought justification for asserting.

<sup>5</sup> It is also clear that Davidson still sees the holistic, open-ended nature of the mental as indicating that FP-explanations take the form, not of deductions from laws, but of the citing of causally efficacious beliefs and desires under the assumption of a suitable background of beliefs and desires (p. 42). This idea, ever-present in Davidson's writings, seems nevertheless to fit rather badly with the admission of generalisations in the paper we are presently discussing. Holism is the topic of section 3.

Davidson also mentions (p. 44) the well-known problem that, for something to qualify as an action, it must be caused by beliefs and desires *in the right way*. However, this does not impact on the question of what kind of explanation an explanation of an action by beliefs and desires is *given* that it *is* an action. It is no part of the view of those who hold that FP involves scientific explanation that explanatory generalisations must exhaust the content of intentional concepts like *action* and *belief*, as we shall indeed see below.



play a role here. In this section, I want to offer considerations that would preempt the need to make such a connection – something that has in any case never been properly done – by suggesting that the intuitive aspects of FP which Davidson (at least as understood by Evidentiary) seems to be appealing to in denying that it is an empirical theory can be accommodated within a view of FP-laws as ordinary empirical statements. Whether or not an alternative framework for understanding FP can be coherently motivated, I would argue that – at least as long as one holds that some kind of generalisations are needed in giving FP-explanations – we do not need to go looking for it just because its principles seem to be things we all naturally know and because they are bound up with the very concepts of belief, desire and action.

Let us first then get clear about what the ‘data’ in this area amount to. (The following observations are due to Evidentiary loc. cit.) It seems clear that we all know, or at least can come to know, that there is an intimate connection between someone desiring, more than anything else, that P, believing that an action of type F will best ensure P and producing an action of type F. Moreover, we can come to know this without doing any research, as we would have to do to find out something of corresponding complexity about purely physical or non-human objects. In addition, it seems reasonable to think that knowledge of such things is in some way constitutive of possessing the concepts *belief*, *desire* and *action*, in a way that knowing that, say, gold has atomic number 79 is not constitutive of possessing the concept *gold*: absent the chemical knowledge I do not lose my grip on what gold is in the way I would seem to lose my grip on what belief or desire or action are were I to fail to grasp the folk psychological ‘platitude’.

Let us first consider the seemingly unlearned status of FP-laws: Does the fact that we do not need to learn these laws suggest that they should be regarded as other than contingent empirical generalisations – as instead ‘a priori’ in some substantive, philosophical sense? In view of much recent research in cognitive science, the answer must surely be: not obviously. Much of this research points to the existence of several innately specified bodies of information: to knowledge of universal grammatical principles, but also, through more recent studies, to knowledge of the fundamental properties of physical bodies, of biological species,

as well as the active principles of the human mind (cf. Pinker 1997: chapter 5 for an overview). Influenced by evolutionary thinking, these modules are viewed as hard-wired in the brain in virtue of the adaptive advantage they have conferred in the process of natural selection. This does not necessarily mean that human infants are born knowing all sorts of interesting facts about the world (though there is some evidence that this is indeed the case – cf. e.g. Spelke 1995) but at least it suggests that knowledge of such facts need not be acquired through learning, rather, it just ‘comes on-line’ at a certain stage in the maturation process (cf. Baron-Cohen 1995 for an elaboration and empirical defence of this kind of position in relation to FP). If FP-laws are amongst those facts with which we are innately endowed in this latter sense (or at least which we can without further ado recognize as a correct systematisation of explanations we are innately able to give),<sup>6</sup> then it will be no surprise that we do not need to do empirical research to acquire them. But then equally this latter point will not support any idea to the effect that FP-laws have a different status from that of other empirical generalisations or facts.

Even if one is sceptical of the idea that FP-laws might be somehow hard-wired in us through processes of natural selection, that is hardly reason to think that their obviousness supports the idea that they are of a different category than ordinary empirical generalisations. If FP is some kind of cultural inheritance, acquired at an early stage in development a similar obviousness would be predicted. Exactly what the mechanisms of this inheritance might be is no doubt unclear, but the very idea of cultural inheritance, taking place at a subliminal level, does not seem implausible.

Given one or the other of these two ideas<sup>7</sup> is on the right track, which there seems no good reason to deny, the unlearned status of FP-laws does not suggest that they are of a

---

<sup>6</sup> This caveat alludes to the possibility that our ability to give FP-explanations may *in the first instance* not involve deriving statements about actions from laws (inter alia), as one who holds a *simulationist* account of our ability to give FP-explanations holds (that is, holds at least this, though she may also hold something stronger). (For the distinction between the so-called theory-theory and simulation-theory accounts of our ability to give FP-explanations, cf. note 16 below.) Note that I am not in this section considering those, whether simulationists or otherwise, who would hold that FP-explanations as such do not involve generalisations at all.

<sup>7</sup> In fact there is also a whole range of intermediate positions, which would see folk psychology as a joint product of innate constraints and cultural inheritance. Such positions could also undergird the overall position I am seeking to defend.

fundamentally different character than, say, the laws of geology or meteorology. In view of the fact that the conceptualist account of FP which Davidson might be seen as offering lacks even a clear expression, we must for the moment conclude that there is no argument here for the idea that FP is in any other business than science.

What about the second feature of FP-laws Evnine draws attention to – that understanding them seems to be essential for understanding the concepts involved, in a way it does not seem necessary to understand the relevant chemical laws to possess uncontentiously empirical concepts like *gold*? This seems to present a more formidable challenge to the idea that the former are empirical generalisations. And indeed the reply I will give is somewhat more intricate.

To begin with it is necessary to underline that the idea that empirical generalisations are completely irrelevant to the identity of the concepts they contain is not very plausible. In particular David Lewis (cf. Lewis 1970, 1972) has presented a philosophical understanding of scientific concepts according to which they are individuated in relation to the theoretical statements they figure in, such that, to really understand a concept like mass as it features in physics, it is necessary to understand and assent to statements such as:

$$F = m \cdot a \qquad \text{and} \qquad F = \frac{GmM}{r^2}$$

Absent knowledge of these statements, a putative grasp of the concept of mass will be undermined. At the same time, few presumably would deny the above laws of Newtonian mechanics an empirical status.<sup>8</sup> The question is thus why one should do so in the case of FP-

---

<sup>8</sup> It is important to clarify at this point that in saying that FP-principles should be viewed as *empirical* laws, I am in no way intending these to be understood as laws which may only be justified by reference to observation. In fact, I take it as given that activity that ordinarily could only be classed as conceptual lies behind the development of many scientific laws. Exactly how the conceptual and the (narrowly) empirical relate to one another is a large topic that we need not dwell on here (nor that of *to what exact extent* particular laws, such as Newton's, might be a priori or conceptual). The burden of my plaint is simply that there is no grounds for drawing a distinction between principles of FP and indisputably scientific laws, which traditionally are known as 'empirical laws'.

## *Is Folk Psychology Different?*

laws, which after all seem to play an essentially parallel role in relation to the possession of concepts like belief, desire and action.

Against such an argument my opponent might point out that a concept such as mass nevertheless has a 'life of its own' in a way those of belief, desire and so on do not. Scientists have over the years changed their views about exactly what mass is; for example most now hold, but did not before Einstein, that it can vary with speed. But in spite of these changes, it is (surely) still mass we are talking about. It does not, by contrast, seem possible for us to find out new things about beliefs and desires, or reject views we once held about them, without in some sense stopping talking about beliefs and desires. Another objection might be that many scientific concepts are ostensibly defined, in the manner first outlined by Putnam 1975, allowing us to latch on to things of the relevant kind, that is, possess the concepts, without knowing much *about* what they refer to. Thus, we can also assume something is of the relevant kind, maybe for several centuries, but be mistaken (*viz. fools' gold*). Again, there would seem to be a grave disanalogy with belief here, which arises if we reflect on the seeming impossibility of a 'fools' believer': someone who behaves as if he has beliefs, but turns out, when we take a closer look, not to. (Evnine *op. cit.*: 13 gestures at these responses.)

But how strong and how significant are these disanalogies really? Consider first the idea that we seem unable to change our theoretical idea about beliefs etc. but not mass. Is that really the case? Developments in twentieth century psychology have plausibly modified FP considerably, with the idea of unconscious and domain-specific intentional states, and with the computer metaphor as a way of understanding how intentional states can have upshots in line with their contents (cf. Fodor 1987: chapter 1). I am not committed to saying that these ideas have actually had an impact on our folk psychological concepts (FP-concepts). Nevertheless, the idea that FP is somehow hermetically sealed from developments in psychology and cognitive science – disciplines which seem, on the face of it, to trade in explanations involving reference to content-involving, intentional states – does not seem very plausible. At the same time, it does not seem very plausible that we will want or be able to

stop talking in terms of beliefs and desires as the causes of action – that we will have to abandon our concepts as FP develops.<sup>9</sup> Of course, this *may* happen, that is, we cannot rule out that a mature cognitive science will have no use for the notions of belief and desire, or even of intentional content; but then the same can be said about *mass* in relation to possible future developments in physics. Either way, there would seem to be no serious disanalogy between FP-laws and FP-concepts, on the one hand, and those of the acknowledged natural sciences on the other. (Note that FP is therefore in one sense not necessarily identical with the empirical theory which we somehow or other naturally know. However, that we are ‘born’ with one particular theory about beliefs etc. does not preclude developing, through scientific research, a theory which preserves the concepts of this first theory in relation to a modified set of truths, any more than theory change in physics implies – as some, notoriously, have concluded – semantic incommensurability.)

Turning to the Putnamian account of the meaning of natural kind terms, one can admit, I think, that the idea of a fool’s believer makes intuitively little sense, whereas that of fool’s gold is relatively clear. We defer to experts about the nature and extent of chemical, biological and physical kinds in a way we don’t about believers or beliefs. However, most philosophers do not these days see Putnam’s (and others’) observations as staking out a radically different theory of reference than that defended by Lewis (cf. e.g. Evans 1973; Crane 1998: 345 ff.). The idea that I can refer to gold whatever facts I believe are true of it is simply not be credited. So, though I can to an extent be wrong about it, and therefore can get corrected by chemists, this does not suggest that I can possess the concept just by bumping into it – by wearing gold earrings or living in South Africa. In addition, I surely get a *better* grasp on the concept of gold – on what gold is – if I learn scientific facts about its constitution, properties and so on. So concept-possession of natural kind terms is by no means completely independent of theoretical knowledge.

---

<sup>9</sup> This does not of course mean that there are not those who would see FP as precisely sealed off from empirical developments in this way, for reasons we shall turn to in sections 2 and 3. The dialectical role of the present point is simply that, absent such further arguments, there is no reason to expect FP to not be influenced by developments in cognitive science and empirical psychology.

## *Is Folk Psychology Different?*

My opponent might nevertheless point out that it is one thing to deny that an understanding of a concept is completely independent of theoretical knowledge, but quite another to assert that this understanding is completely dependent on such knowledge – implying that both hold for FP-concepts, but only the first for concepts like *gold* and *water*. However, we have already seen that it seems reasonable to suppose that FP-concepts would survive some changes to FP itself, so it seems that this response must be wrong. What we are left with is the intuition about the difference between the possibility fools’ believers and fools’ gold. However, it is relatively straightforward to dilute even this. It is a platitude that we can be wrong about particular attributions of beliefs and desires, and of predictions of action. If certain subsets of the totality of such mistaken attributions are based on the use of mistaken parts of our folk psychological theory, we can presumably see these as fools’ *beliefs*, desires and so on. Indeed, not even the idea of a fools’ *believer* seems ruled out in principle. It seems conceivable we should want to withhold attributions of beliefs and desires completely to a being which behaved in ways which otherwise would warrant such attributions, for example, if it turned out that the being possessed an internal architecture which did not seem capable of supporting conscious thought (cf. Clark 1993: chapter 10).

Moving to discussions of this kind – involving bizarre thought experiments to establish a principled point – are, nevertheless, inveterately unsatisfactory. A more important and deeper reason for disregarding the apparent disanalogy between fools’ gold and fools’ believers would thus seem to be that it seems plausible to see its basis as lying in the different epistemic relations we bear, and which it is general knowledge we bear (albeit perhaps at a tacit level), to the properties of gold, on the one hand, and those of thinking human beings on the other. We have some kind of natural access to FP, and everyone knows this about everyone else (tacitly). That is, we take it that everyone, including ourselves, has an understanding of the basic inter-relationships between beliefs, desires and action, and that we classify things as believers according to whether they satisfy these basic truths (which we likewise assume are truths). We have by contrast perhaps no natural access to the principles and laws of chemistry, nor do we expect anyone except professional chemists to have this.

Now, the idea of a fools' something-or-other – a fools' F – rests most fundamentally on the idea that one can systematically misidentify something as an F; that is, one can systematically claim or think that certain things are of kind F, even though this is false (or the converse), and yet not be in danger of losing one's grasp on the concept F. It is this that seems unthinkable, or at least very difficult to imagine, in relation to the concept *believer*.<sup>10</sup> Often in the literature this difference is attributed to a difference in the concepts, and it is suggested that *believer* must be somehow unempirical. My suggestion is instead that the difference can be explained by differences with respect to what people know and know others know. To deny that someone really understands what a belief or believer is when she fails to attribute a normally functioning human being with beliefs, or attributes them to a rock, but to not do something similar when someone fails to identify samples of gold as such, or calls something gold that is not gold, is reasonable given this difference in known access to the relevant theories. It seems that whether one can be said to possess a concept or not can, in certain contexts, be a matter of what is generally expected of one by one's peers, rather than a purely objective matter of what one knows about a kind of thing. I believe this also fits well with the fact that concept-possession can come in degrees: at which point in the progression of knowledge about a kind of thing one can be said to actually possess the concept of it is ultimately somewhat arbitrary, and thus is susceptible to influence by factors of the kind I have just mentioned. Or perhaps what we mean when we talk of fools' gold is that we can *kind of* possess the concept of gold even though we know very little about it – at the same time as one possesses it in a much more definite sense after completing a chemistry degree. We don't tend to say the same things in relation to the concept *belief* or *believer*, but my suggestion is that this is not diagnostic of some deep difference between the concepts, but rather results from the fact that everyone tends to assume that everyone knows everything about what beliefs and believers are. Thus, in the context of a chemistry lab, it seems it would be reasonable for the researchers to question

---

<sup>10</sup> It might be objected (as an anonymous referee did object) that this is not unthinkable, even without resorting to bizarre thought experiments. For non-human animals are often classified as believers and yet might turn out not to be, if it turns out that language is essential for thought (or conversely). Though various responses might be proffered to this point, I will pursue my own line independently of it, not least because it seems at least intuitively reasonable to view FP as genuinely, i.e. non-metaphorically, applicable only to humans.

a newcomer's conceptual ability were she to manifest an absence of chemical knowledge of gold – to say 'She doesn't seem to know what gold is' – even though they wouldn't say something similar about ordinary people who lacked the same chemical knowledge. Similarly, the idea of fools' gold makes little sense within the context of a chemistry lab. My suggestion is that our ordinary situation with respect to folk psychology is a bit like that of the chemistry lab with respect to gold. If that is right, the differences Evidine notes between our attitudes to the concepts of belief and of gold do not have any deep consequences for the kind of generalisation they figure in.

In this section I have been concerned primarily with Simon Evidine's exposition of the thoughts behind Davidson's anomalist view of psychology. Nevertheless, I think this exposition, though elementary, is important, and articulates for many the common-sensical motivation for Davidson's view. I hope therefore to have shown that this basis is susceptible to reinterpretation in view of recent developments in cognitive science, and of sober reflection on the nature of concepts. Of course, many would want to say much more about all this, especially reference and concepts, but my aim here has been to steer a course that as far as possible does not impinge on big philosophical controversies, at least as things currently stand in the debate; and in so doing, to indicate why there is no reason to regard FP as other than an (albeit no doubt low grade) empirical theory.

## *2. Normative explanation*

The previous section was aimed at showing that there is no reason to see the seemingly a priori and in some sense conceptual status of FP as indicating that it is other than an empirical theory. In this section, we will be concerned with the idea that FP gives normative explanations. So far, we have seen that, insofar as one acknowledges the role of generalisations in FP-explanations (as Davidson does), there is no obvious prior motivation for introducing normativity to explain how these generalisations can be conceptually true, for it is far from obvious that we should view them thus. Even so, I haven't shown that FP-



generalisations aren't conceptual, so if there turn out to be good arguments for viewing FP-explanations as normative, and this entails that any FP-generalisations are be conceptually true – or even explains what it is for them to be thus – this would be consistent with what I have so far argued. In this section, without assuming anything about a relationship between the putatively conceptual and normative aspects of FP-explanation, I will adduce considerations that militate against the coherency of the view that it is normative.

As noted in section 1, Davidson's remarks seems to open for the idea that it is the normative character of FP-explanations that explain why FP-generalisations cannot be seen as ordinary empirical laws. Further, he stresses that norms of FP are special – and not to be viewed as on a par with the constitutive principles of the various sciences – because they are *ours*: 'reason-explanations [i.e. FP-explanations] make others intelligible to us only to the extent that we can recognize something like our own reasoning powers at work' (p. 47). I will return to this reflexive aspect of FP-explanations and its putative connections with normativity below.

First, however, I want to introduce the views of McDowell, who has been amongst the most influential of those who have taken Davidson's views on anomalism seriously. In the course of several papers on diverse topics, McDowell stresses that what we are calling FP-explanations have a radically different character from those of science. The following two citations are representative:

'Intentional explanation makes an action unsurprising, not as an instance of the way the world works (though of course it does not follow that an action is not that), but as something which the agent can be understood to have seen some point in going in for.'  
(1980: 126)

'[P]ropositional attitudes have their proper home in explanations of a special sort: explanations in which things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be. This is to be contrasted with a

style of explanation in which one makes things intelligible by representing their coming into being as a particular instance of how things generally tend to happen.’ (1985: 389)

But what is McDowell really after here? And is it coherent? I want to begin by taking what these quotes say pretty much at face value, before moving on to what seems to be the driving thought behind McDowell’s (though only I think to an extent Davidson’s) insistence that FP-explanations are different from scientific ones.

Consider first the second citation. McDowell says that when we explain actions and beliefs, we are making them intelligible by showing that they conform to what rationally ought to be the case. This is contrasted with making something intelligible by, roughly, subsuming it under a general law. Factoring out the talk of intelligibility, this comes down to saying that FP-explanations explain a belief or an action by revealing it to be as it rationally should be.

On the face of it, however, this suggestion does not seem coherent. Let us imagine X comes to believe that B on the basis of believing A and if A, B. McDowell’s suggestion seems to be this:

(1) X believes A

X believes that if A, B

If one believes A and if A, B, then one should (rationally) believe B

X believes B

The problem is just that the third ‘premise’ seems to be redundant.<sup>11</sup> How can the fact that someone *should* do something – in and of itself – help to explain *the fact that* they do something? Note my objection is not that the explanation would only work were we to add something about it generally, i.e. empirically, being the case that people believe B when they

---

<sup>11</sup> Note that I don’t mean to suggest, in presenting the putative explanations in this section in this way, that explanations necessarily take the form of arguments. The format used is merely a convenient way of contrasting different interpretations of McDowell’s thoughts.

believe A and if A, B; or that they generally do what it is rational to do (though I am not denying this either).<sup>12</sup> Perhaps the explanation is fine just as long as one deletes the normative statement. My point is simply a registration of confoundment at the idea that such a statement can be independently relevant in an explanation.<sup>13</sup>

But is (1) really what McDowell intends? The first quotation would perhaps suggest something different: ‘... intentional explanation makes an action unsurprising [...] as one which the agent can be understood as *having seen some point in going in for*’ (emphasis added). This suggests that the right form of the explanation of X’s belief in B might be the following

(2) X believes A

X believes if A, B

X believes that if one believes A and if A, B, then one should believe B<sup>14</sup>

X believes B

This is intuitively much better than (1). Moreover, it gets across the idea that when one forms a belief or acts rationally, one sees oneself as doing what is rationally required of one. However, the question arises now as to whether (2) is any way special as a kind of explanation.

In (2) we take as part of the explanation that X sees her belief-B-formation as in some sense ‘the thing to do’. But now how do norms figure in this explanation? One might say that X has a belief about a norm of inference, about what the relevant norms are: she has a belief about what one *should* do in certain circumstances. But clearly this is not the same as making

---

<sup>12</sup> This was Hempel’s (1965) objection to the idea of purely ‘rational’ explanations. Ostensibly normative explanations aside, both Hempel and McDowell assume that subsumptive explanations exhaust the realm of genuine explanations, an assumption I do not make use of in my arguments.

<sup>13</sup> Note also then that it should not be understood as begging any fundamental questions about the fact-value distinction (see also below in the text).

<sup>14</sup> X needn’t of course believe what it is stated she believes in precisely this way. One way of this statement being true would, for example, be for X to be consciously aware of the inference from A and if A, B to B itself being, as Peacocke (1991) puts it, *primitively compelling*. The clause is meant simply to get across the idea that X herself sees what she comes to do as warranted given her prior beliefs.

something ‘intelligible by being revealed to be [...] as [it] rationally ought to be’. If X had some other, *false* belief about how things rationally ought to be, then a similar explanation of why she formed some other belief, C, would still work, viz:

- (3) X believes A  
X believes if A, B  
X believes that if one believes A and if A, B, one should believe C  
X believes C

Here we can hardly say that the explanandum is revealed to be as it rationally ought to be, even though it is – plausibly – explained.<sup>15</sup>

But perhaps I am over-complicating things. Surely all we need to register to explain X’s belief in B are the two beliefs whose contents entail B:

- (4) X believes A  
X believes if A, B  
X believes B

McDowell might simply point out that this is an intuitively good explanation, and that, insofar as it does not make use of laws, the insight it gives must have another source; and that the idea of what is normatively correct springs naturally to mind.

However, if what is normatively correct is what ‘makes’ the explanation, then rendering (4) as (1) should be possible; but (1) seems incoherent. Further, it is not clear why the insight offered in (4) needs to have some special source. Compare (5):

---

<sup>15</sup> Essentially this line of thinking is also to be found in Henderson (1993, cf. especially chs. 5 and 6): I take my point to extract the intuitive essence of Henderson’s case. Much of Henderson’s book contains discussion of points relevant to the present paper, but his stress is on the social sciences and the relationship between interpretation and explanation, whereas mine is on the need and very coherence of a special kind of explanation. See also the Conclusion for brief discussion of the relationship between interpretation and explanation in relation to folk psychology.

(5) The Titanic crashed with an ice-berg

The Titanic sank

This seems like a reasonable explanation as it stands, at least as good as (4), but there is presumably no question of seeing the sinking as something that ought, rationally, to have happened, given the preceding event. One might protest that a law connecting the cause and effect in (5) must be being tacitly assumed, or at least being assumed to be available, in order for the explanation to appear convincing; whereas this is not the case with (4). But saying that, in the present dialectical context, would simply beg the question. Maybe we can also give an empirical law to subsume the beliefs and their interrelations; at least, we have seen no coherent reason so far for taking the prospects for finding laws to be less remote in (5) than in (4).

In the following section, we will consider the holism of the mental, which for many has signalled the principled futility of seeking even rough or *ceteris paribus* intentional psychological laws. Nevertheless, McDowell does not rest his case on holism, and, insofar as he is of course aware of it, would appear to see the normative nature of psychological explanation as its fundamental distinguishing feature. Yet we have seen that the normative conception of FP-explanation is not only mysterious but unmotivated.

Having said that, I do not doubt that a McDowell-sympathiser would regard this dismissal as unwarranted, so let us dwell for a while on what she might say in response to my arguments. To begin with, it might be argued that I am simply begging the question. In particular, one might press the point that McDowell's views on explanation presuppose that there is no sharp divide between the normative and the descriptive, between fact and value; whereas my objections exploit precisely such a divide. However, this would not be fair appraisal of the dialectical situation. As I see it, McDowell's overall conception of psychology and man's place in nature (for more on which, see below) rests, at least in part, on the *intuitive* plausibility of the idea of normative explanation in psychology, which is taken as

a kind of datum. Now what is intuitive can no doubt vary from person to person. However, it strikes me that insofar as intuitions are relevant to philosophical discussions, they should be at least amenable to being consistently and coherently spelled out. The two quotes we have given from McDowell do not even seem to express the same idea. Moreover, given that neither seems to present a coherent notion of a special kind of explanation – something which seems to be as obviously the case as anything could be – it seems the intuitive support that McDowell seeks in the nature of FP-explanation simply does not exist.

But perhaps the context of my appraisal of McDowell's idea is too narrow. For example, he insists that FP-explanation is intimately and inextricably bound up with evaluative practices of criticizing and endorsing others' beliefs and actions, including one's own (cf. e.g. 1985: 392). These practices are viewed by McDowell as dynamic and irreducibly social in nature, at the same time as they are also reflected in the nature of conscious deliberation, which precisely takes the form of seeing oneself as subject to requirements that are independent of the actual decisions one makes. When one predicts that someone will form a certain belief, one is at the same time standing prepared to criticize her if she doesn't, or endorse what she says if she does. At the same time, one stands open to correction from others as to whether what one lays down as the norm is in fact correct. There are no external givens as to what one ought to do or believe – except perhaps the nature of the external world. Perhaps, then, FP-explanations are of a different character from ordinary empirical ones because they must somehow reflect this inherent but ineffable *being-subject-to-a-norm* aspect of deliberation. FP-explanation, in seeking to track the actual contours of individuals' thought processes, is constitutively tied up with justification of such thought process.

One sympathetic to phenomenological methods in philosophy might feel a sense of accord with this way of describing the circumstances. To me, however, it too lapses into incoherence. How can I really be seeking to do the 'right' thing if, ultimately, there is no standard, independent of what anyone actually does in particular cases, to constitute it? What does it mean to speak of standards under such a description? Moreover, even if we can

somehow make sense of McDowell's dynamic and societal picture of standards for rational action and belief-formation, it fits badly with the idea of FP-*explanations*, insofar as these seek to give an account of *particular* episodes of thought and action. In relation to these, what one should, objectively, do or believe – as opposed to what any individual thinks she should do or believe – seems, again, on the face of it irrelevant to understanding what any particular agent at any particular time did or came to believe.

What about the idea of reflexivity – that FP-explanations somehow make sense in terms of *our own* norms, as opposed to the norms – i.e. laws – that govern aspects of physical reality? This seems to be important for McDowell, as well as for Davidson as we have noted. But it is hard to see how it can mark out FP-explanation as special in the way McDowell and Davidson want it to be. Note first that insofar as all schemes of understanding in some sense are ours, all laws or norms could be said to be our own. The question is in what way norms in FP-explanation are in some more fundamental sense ours; ours in a sense other than that they describe certain de facto regularities in *human* behaviour.

Perhaps the idea is that FP-explanations are given relative to one's own imaginative capabilities in relation to others, and to the output that spontaneously springs to mind when one applies these capabilities. In other words, explaining why X came to believe B involves putting oneself in X's situation and considering what might have brought oneself to believe B. This idea is reminiscent of the so-called *simulation theory* of FP-explanation as recently discussed in cognitive science.<sup>16</sup> However, the notion of reflexivity involved in ST does not seem to correspond to that which McDowell (or Davidson) have in mind. The latter is not

---

<sup>16</sup> Here the focus has been on the fact that all ordinary humans can give FP-explanations without the conscious use of any kind of theory, or any kind of explicit instruction in childhood. Cognitive scientists have been concerned with the cognitive basis for this ability, and two broad proposals have emerged: the so-called theory-theory (TT) on the one hand, and the so-called simulation theory (ST) on the other. According to TT our ability to give FP-explanations is based on a body of implicit or tacit knowledge of a theory of empirical generalisations (perhaps of the kind outlined by Churchland op. cit.), which we are not generally aware of. Our ability to give FP-explanations is thus to be viewed as in many ways on a par with our ability to utter and comprehend structured sentences of a natural language, as construed by Noam Chomsky and his followers: both involve a domain-specific body of unconscious knowledge. According to ST, our FP-ability is based instead on the use of our own decision-making and belief-forming mechanisms which we use 'off-line' to generate predictions and explanations by feeding them, via a process of imaginative projection, with the beliefs and desires we take the object of our explanations to occupy (cf. e.g. Gordon 1986). In the early stages of the debate, the ST versus TT question was viewed as amenable to empirical adjudication, but most recently there has been a deal of attention focused on understanding what the distinction itself is meant to amount to (cf. Stone & Davies 1996).

just a first-person reflexivity, since I stand subject to others' criticisms and conceptions of the norms in my own reasoning. Connectedly, the explainer's conception of the interrelations between the explainee's beliefs, desires and actions, is not, on a simulationist understanding, in the first instance normative: I do not, according to simulationists, see X's inference to B as something she *ought* to do, but merely something I would have done in similar circumstances. The explanation is intelligible in relation to my own intuitive grasp of the situation, and the fact that I am myself involved in such situations, not in relation to the way things rationally ought to be. (I will return to ST and its independent impact on our question in section 3.)

A further possible interpretation of the idea of reflexivity is that it reflects the way in which certain inferential rules are a *transcendental precondition* for any meaningful enquiry, inference or set of beliefs, and thereby – in precisely some special way – ours. The rules of logic are perhaps the most obvious examples of such rules. However, even in relation to logic, viewing rational rules as transcendently justified – even the idea that this has a clear meaning – is controversial. And in any case, FP is, as it is generally construed, concerned with much more than explaining episodes of making inferences or performing actions that can be licensed deductively. (Note also that, whilst it may be that there is some kind of conceptual connection between the explanans and explanandum in FP-explanations, this is not, as we saw in the previous section, a reason to deny the explanations are based on the same kinds of generalisations we find in the rest of science.)

In his most recent writings, McDowell has sought to re-introduce the Aristotelian idea of second nature as a way of understanding the place of mind in natural world, and to close this section we will briefly consider its impact on our discussion. According to McDowell (1994: lecture 4), there is a tract of reality which is our peculiarly human, rational nature. This involves a realising of natural capacities that otherwise are amenable to a nomological, scientific understanding, but whose realisation in us – as opposed to its realisation in 'dumb' animals – is not. This is our second nature. It involves those aspects of our human nature that might properly be regarded as cultural (in a non-relativistic sense), including all those that depend on the use of language and ethical thinking. What is arresting about this idea is that



McDowell is committing himself to a genuinely *ontological* category bearing the characteristic qualities of human culture and language-use, rather than the latter being either reducible to something essentially non-cultural, or else being seen as a ‘projection’ onto an otherwise wholly physical reality. Moreover, it is the realm of our second nature to which reason explanation, that is, FP-explanation, is proprietary. Given that second nature marks itself so radically off, ontologically, from the realm of nomological nature, it is *ipso facto* that FP-explanations won’t be scientific or law-like.

McDowell sees no obvious barrier to this way of viewing our natures – as partly ‘re-enchanted’, as he puts it – at the same time as he sees lots of advantages (in relation, in particular, to the problem of explaining how empirical content and knowledge is possible). However, a driving thought behind what he sees as the advantages is precisely that there is a special kind of understanding proprietary to the realm of reason. In relation to the present dialectical context, then, one should not take the move to second nature as being in some way independently motivated. Rather, our question must be whether it does anything to make this commitment, which is what is *de jure* for us, more lucid or intelligible.

I cannot see that it does. Second nature is meant to be ontologically distinct from first nature.<sup>17</sup> But – given this is not merely a question of different and conflicting perspectives on one and the same nature – then it seems we have merely two distinct (though related) tracts of reality with their own characteristic modes of understanding (i.e. of being understood). What surely does not follow, without further argument, is that this can motivate or defend the idea that FP-explanations are different in kind, fundamentally, from scientific explanations.

---

<sup>17</sup> Some might feel the ontological conception of second nature is an exaggeration. Might it not instead simply be that FP offers precisely a *perspective* on ourselves that is somehow illuminating; a perspective from which it makes perfect sense to see the explanations of action we give, on the one hand, and our evaluations of these, on the other, as two sides of the same enterprise – a perspective existing alongside that in which explanation is, by contrast, essentially just a matter of describing patterns that are objectively there? However, it seems there is something fundamentally unsatisfactory about making a distinction of this kind. Ordinarily, when there are two conflicting perspectives on some issue, at most one reflects the true nature of the case. Of course, the idea of *perspective* invoked in the response we are considering is meant to be metaphorical, but it is unclear how this is meant to help since we can presumably only rely on what perspectives generally are, in ordinary parlance, to understand the metaphor. It might also be the case that, in talking of perspectives, an association is being sought to the idea that reason-explanation is a matter of *seeing* others and oneself *as* subject to norms. But it seems that, absent an understanding of the first, metaphorical kind of perspective, this can only amount to saying that we are not really subject to the relevant norms, *or* that we are and that they thereby characterise a genuine but empirically describable tendency in us.

Biological explanations are scientific, but they concern for many the ‘biological world’, rather than the ‘physical world’. If there is a world proprietary to the psychological, why should an understanding of this be ‘unscientific’? Indeed, what would it *take* for the understanding of this world to be in this way different in kind? One owes some positive account as to what distinguishes FP-explanations as special, and just appealing to the idea of second nature does nothing to establish this. That there may be limits to our scientific understanding, especially of the mind, is not what is at issue here. There may indeed be such limits, though as things stand there seems no reason for denying that FP offers *some* kind of insight into human action and thought processes. My point is that McDowell presents no way of understanding this that cannot be reconciled with models of ordinary scientific explanation.

McDowell and his sympathisers will no doubt still insist on the possibility of a third option – between scientific explanation and no explanation at all. I have tried to show in this section how difficult it is to give this option any coherent content in relation to the idea of normativity.

### *3. Holism*

In this final main section, we will consider what impact the allegedly holistic aspect of our reasoning has on the question of what kind of explanation FP-explanations offer.

As noted in section 1 (note 5), holism still seems to be one of Davidson’s reasons for denying an empirical status to intentional generalisations. In the past, he has quite directly connected the holism of the mental – and indeed, his anomalism generally – to Quine’s doctrine of indeterminacy in relation to radical translation (cf. Quine 1960; chapter 2; Davidson 1980: 222). In recent years, Davidson has distanced himself from this connection, perhaps partly because the idea of indeterminacy for Quine is intimately involved with a sceptical attitude towards the reality of intentional states, which Davidson does not (straightforwardly) share. We need not, however, concern ourselves with such metaphysical questions here. The important question for us is whether and in what way the holism of the

mental – whatever its connections to the indeterminacy of translation and mental realism – marks out FP-explanations as special in relation to other kinds of explanation.<sup>18</sup>

The relevant observation for us is thus simply that the reasoning we undertake that leads to action and belief- and desire-formation can seemingly only properly be understood in relation to a whole host of propositional attitudes, never just some neatly demarcated subset of those an individual can be said to hold. If you explain why some person crossed the road at a zebra-crossing by saying that she wanted to get safely to the other side, then you also assume, possibly tacitly, that she knows what a zebra-crossing is meant for, that she believes car-drivers will stop at such crossings, that car-drivers have eyes and can see, that cars can be dangerous because they are built of metal and move fast, that metal is heavy, that heavy, fast-moving things generally do damage to objects they hit, that people generally seek to avoid damage to their person, etc., etc. – and without any obvious limit. There seems no point at which one can stop and say ‘*These* are the relevant attitudes’, since it is (or seems to be) always possible to imagine other beliefs and desires that the protagonist might have that would undermine the intuitive reasonableness of what she does or believes.

This holistic aspect of our reasoning<sup>19</sup> has been an obstacle to understanding human processes of reasoning in cognitive science, especially AI, something which in its turn has been taken as indicating a principled, epistemic barrier to a computational and/or law-like understanding of the processes of belief-formation and decision-making (cf. Fodor 1983). But others have gone further and seen it as indicating that FP-explanations must be of a principledly different character than those we find in cognitive science – or even that FP-explanations could never be seen as scientific explanations at all. A recent example is Pickering & Chater (1995), whose views I shall comment on below.

Before that it is worth considering more generally just how and why holism is meant to give support to the idea that FP-explanations are different from other scientific explanations. One reason for thinking this might be that holism is intimately bound up with

---

<sup>18</sup> Cf. also the remarks on interpretation in the Conclusion.

<sup>19</sup> Which, for the purposes of this essay, I will simply accept. For a relevant challenge, cf. Botterill & Carruthers (1999: 71 ff.).

the normative and dynamic aspect of our understanding others, as stressed by McDowell. The beliefs and desires which can be said to underlie a given inference cannot, given holism, be specified determinately once and for all: however specific one tries to be about why our protagonist crossed the road, one will never have a logically deductive line of reasoning, and one will always be in a position where alternative lines of action might have seemed reasonable given that certain background assumptions, as yet unmentioned, were entertained. In view of this, it might seem reasonable to say that when one attributes someone with a given set of beliefs and desires to explain an action, one is simultaneously committing oneself to saying that these attitudes make the action intelligible in light of what (oneself believes) one rationally *ought* to do, a judgement that in its turn can itself only be understood in relation to a dynamic social practice of mutual criticism and evaluation.

However, it is not clear that we can make any sense of this line of thought, as we saw in the previous section on McDowell. Moreover, it is far from clear that holism demands it. Let us assume for the sake of argument that FP-explanations do tend to have the kind of ‘open’ character described above. This doesn’t seem to prevent one giving explanations of particular actions and beliefs that are in and of themselves satisfying, even though they always can be fleshed out further. Moreover, it seems many non-psychological explanations are also in this way holistic. Take again example (5):

(5) The Titanic crashed with an ice-berg

The Titanic sank

Even though this is intuitively satisfying as an explanation, we can fill it out seemingly indefinitely, casting ever more light on the causal process that eventuated in the Titanic sinking. We could mention the size, shape, temperature and material constitution of the boat and the ice-berg; the speed at which they were moving when they collided; the angle at which they collided; the speed with which the water entered the vessel after collision; and the temperature of the water and its natural dispositions. Is that all? It is not: in principle we

would also need to talk about the prevailing weather conditions, about precisely what was done to prevent the water entering the boat, the realised and unrealised efforts made to salvage the boat, and so on, without end.

Now, someone might still want to claim that it is reasonable to suppose there is a low-grade, *ceteris paribus*, but nevertheless genuine law connecting boats striking ice-bergs and subsequently sinking. The morass of detail I have alluded to as part of a full explanation of any particular sinking would then be covered or assumed in the *ceteris paribus* clauses. I don't need to take a stand on whether this is correct. For us the question is: Is saying there is a law covering boats hitting ice-bergs and sinking any *more* plausible than saying something similar for particular kinds of beliefs, desires and actions? Note that in both cases, the assumptions about what *ceteris* must be *paribus* for a putative law to operate will refer to events and properties *at the same level of description*, in other words, not just to the principles of lower level mechanisms that must be assumed to be operative, but to same-level processes and events that might or might not intervene. I submit that there is no case for saying that there is any more meaning or sense in invoking laws to subsume phenomena like boats hitting ice-bergs and sinking, or stones hitting windows and the latter breaking, than invoking laws to subsume events like believing it is raining, desiring not to get wet and taking out one's umbrella or rain-coat. Perhaps it is reasonable to assume access to very rough generalisations about the behaviour of different kinds of everyday objects in relation to one another – for example, that hardish, heavyish objects tend to break fragile things when they come into contact with them. But then equally one can say that beliefs about what is best to do in certain circumstances to obtain a desired good tend to lead to people performing that action.

An alternative way of putting this point is that there is no case for taking explanations of particular physical events to be any less *idiographic* than explanations of particular mental events. An *idiographic* explanation is one in which light is cast on an event or phenomenon, not by relating it to general tendencies or laws, but by giving detailed information about the precise particular circumstances surrounding it and leading up to it. Since this rich morass of information will be peculiar to the event in question, one will not be able to credit the idea of

a general law as functioning in its explanation, for the exact combination of circumstances in question will be in effect unrepeatable. History is often viewed as giving such explanations, and according to e.g. Colin McGinn (McGinn 1979: 39) something similar holds for FP-explanations: we explain actions and the formation of propositional attitudes by seeking ever more detailed information about the particular beliefs and desires entertained by a given subject. Now, it is controversial whether one can make sense of explanation that is through and through idiographic. My point is that it seems to be apt to the same degree in relation to physical as to psychological happenings.

In response to this, someone might admit that (certain) physical events can be explained idiographically, like mental events, but insist that the former at least *can* also, in principle, be exhaustively explained by reference to fundamental physical laws, together with some total specification of a set of relevant physical initial conditions. Note first that this claim rests on pure metaphysical speculation, on some kind of faith in the completeness of physical explanation that has never been demonstrated. But then whether the claim is correct or not is beside our point. For, if we accept the speculation on which it is based, why shouldn't actions and mental events also be amenable – in principle – to explanations by exceptionless laws (perhaps some refined descendent of decision theory together with a theory of how it is physically instantiated in us), plus sets of initial conditions (the particular beliefs and desires)? Without appealing to one or other of the ideas discussed but discredited in the foregoing, we seem again unable to draw a principled divide between the mental and other domains.

Are there other ways in which holism might suggest that FP-explanation enjoys a fundamentally different status from other kinds of explanation? An idea which Pickering & Chater (op. cit.: 317) suggest is that holism shows that we rely on some shared but unconscious and perhaps inarticulate set of background of assumptions when giving FP-explanations, against which drawing attention to one or two particular beliefs and desires to explain others or an action can be illuminating. However, the obvious problem with saying this is that it seems it might again well apply to most non-psychological explanations we give

– the sinking of the Titanic or the breaking of a window. Is there any principled difference between the use of background knowledge in giving explanations of the latter kind, and FP-explanations?

According to many cognitive scientists, the tacit background used in giving explanations can be seen, at least in part, as a kind of internalised theory about some portion of the natural world – physical objects, biological creatures, people – that is largely determined by innate precursors (cf. section 1). If that is correct, there will be rough generalisations standing behind our intuitive explanatory practices both in relation to physical and psychological explanations.

According to others, the relevant background at least when it comes to FP is not any kind of theory, but rather some kind of skill we possess to generate explanations. Here we return to the idea of simulation theory touched on in section 2. According to this, we generate FP-explanations and predictions by using our own decision and inference mechanisms ‘off-line’, but without the help of any kind of theory (cf. note 16). Of course, something similar might be the case with non-psychological explanations – that we somehow apply our imagination to new cases to generate predictions and explanations without relying, even tacitly, on anything like a theory. Though this seems eminently possible, concrete research on the matter has not as far as I am aware been pursued.

Does this mean that were the empirical research in cognitive science to reveal that FP-explanation is based on something like simulation, whereas our physical explanations are based on something like a theory, that FP-explanation would after all have been shown to be different – non-scientific? I think this would be a rather dubious conclusion to draw. Theories in cognitive science about the basis of our abilities to give certain kinds of explanations concern first and foremost that: our *abilities* to give explanations of certain kinds. They do not therefore impact directly on the constitutive nature of the explanations thus given.<sup>20</sup> That we naturally do not use generalisations in giving FP-explanations would not imply that such

---

<sup>20</sup> I thus assume (the standard view) that explanations are first and foremost abstract informational products, rather than being essentially tied to the process of giving them (cf. e.g. Ruben 1995: 8-9 for a defence of this idea).

generalisations do not exist and could, or ideally should, be given; nor does the fact that generalisations are used imply they must or should be. Indeed, in exploiting one's own reasoning mechanisms in order to give such explanations, it seems one is relying on a certain degree of regularity in the operation of such mechanisms of a kind that is presumably amenable to capture by some kind of empirical generalisation.<sup>21</sup> It is in general hard to see how – given that we possess distinct abilities for giving different kinds of explanation – some kind of regularity is not being exploited, be it in the form of unconscious knowledge of an empirical law, or simply the regular operation of some part of the cognitive system. If that is so, it is unclear that the kind of *explanation* generated by a simulation module should be expected to be essentially different from that generated by a theory module.

I take the foregoing considerations to suggest that the holism of the mental does not show that FP-explanations are of a different character from ordinary scientific explanations. As far as the dialectical thrust of this essay is concerned, that is all I am interested in establishing. A further issue – which in many ways is the main topic of Pickering and Chater's article, and will also illuminate our discussion from a slightly different angle – is to what extent scientific psychology, or cognitive science, is at odds with FP: to what extent cognitive science and FP constitute mutually exclusive domains of understanding. Pickering & Chater argue they are: cognitive science gives computational explanations of either encapsulated capacities or structural features of the mind, whilst FP employs a form of holistic, 'content-sensitive' form of explanation. In a reply to this article, Morris & Richardson 1995 argue for the opposite view, adducing examples suggesting that both cognitive science and FP have broader explanatory ambits than Pickering & Chater suggest they have, thus blurring any proposed divide between them. We need not go into the details of this debate here (though my sympathies lie squarely with Morris & Richardson). It is diagnostic of the dialectic that is relevant to us that in a response to Morris & Richardson (Chater & Pickering 1997), Pickering & Chater argue that debating whether extant cognitive

---

<sup>21</sup> On some theories about the nature of tacit knowledge, the very fact of this causal regularity would amount to possessing tacit knowledge of the relevant generalisations: cf. Davies 1994.



science and what goes under the name of folk psychology in fact overlap is irrelevant to the question of whether what they call ‘explaining by justification’, which applies in FP, and what they call ‘explaining by causes’, which applies elsewhere, are reconcilable; and that, as far as this latter question goes, the answer is that they are not. However, Pickering & Chater do practically nothing to undergird this claim – presupposing, one takes it, the authority of thinkers like McDowell and Davidson on the matter. Yet, as we have seen in this paper, there seems to be no coherent set of ideas to back up this authority.

### *Conclusion*

My aim in this piece has been to examine and to seek to explode a certain, in some circles, received view of the nature of intentional or folk psychological explanation, that is, the scheme of explanation involving, centrally, the notions of belief, desire and action. I have been arguing that none of what I take to be the three main ideas that are commonly taken to ground this view provide a coherent such ground: the idea that folk psychological generalisations are conceptually true, the idea that FP-explanations are somehow normative, and the idea that they are holistic.

In concluding, I want to make four further points of a mainly clarificatory nature.

Firstly, it would be a misunderstanding to think that I have been attacking straw men by taking each of the above ideas separately – that if only they were seen in their proper context in relation to one another, the true force of the idea that FP-explanations are other than scientific would clearly emerge. For, as the reader will have noted, I have not insisted on holding these aspects apart from another, on the contrary, I have sketched several possible lines of connection between them. My arguments have of course been directed at the individual ideas, rather than the whole package. But this seems eminently reasonable as a strategy in undermining the overall view they are meant to support, for what I have been demanding is nothing more than a clear statement of which aspect or aspects are meant to ground the idea that FP-explanations are different. None such has been forthcoming.

Secondly, and connectedly, one might feel that the case against those who propose that FP-explanations are different should also take on board the tight connections between interpretation and explanation that feature in at least Davidson's work. To this I reply that nothing of what I have said conflicts with the idea that when *interpreting* someone, we must find them more or less rational, as Davidson's principle of charity requires; or with the idea that interpretation is indeterminate. What I have argued entails that, if either of these ideas are right, they cannot coherently imply that there is some special kind of FP-explanation, distinct from other 'scientific' explanations. This commitment doesn't mean either that people generally are not optimally rational; it just means that when it comes to explanation, we are concerned with actual antecedents of behaviour, not norms.<sup>22</sup>

Thirdly, some readers may feel that I have neglected the putative world-involving nature of propositional attitudes, and the impact this might have on the question of explanations involving them. Now it is of course a large issue in philosophy of mind whether and to what extent intentional states should be characterised by reference to, or even be seen as partially constituted by, aspects of thinkers' environments; and it is also a large issue whether and in what way this has bearings on the viability of a scientific intentional psychology. In view of this, one might perhaps think that the world-involving nature of beliefs and desires could give support to the idea that FP employs a different style of explanation than that we find in science generally. However, it is a fact that many who hold that intentional states are in some sense world-involving also believe in scientific intentional psychology (e.g. Fodor, Burge). Moreover, it strikes me that the only way in which one would think that world-involvingness suggests that FP is other than scientific would be if the nature of the world were seen as some kind of *norm* for the formation of beliefs ('one should believe the truth'). But then the proposal would lapse for essentially the same reasons adduced in section 2, where we tried in vain to find some coherent core to McDowell's idea of FP-explanation as a based in rational evaluation and criticism.

---

<sup>22</sup> Henderson (1993: chs. 2 and 3) argues for the further idea that good interpretations should utilise, not a principle of charity, but a principle of *explicability*, which enjoins finding what people do explicable. Hence interpretation is bound up with explanation. This is a further commitment that does not concern me here.

Finally, it may have struck some readers that I have ignored a further idea that is often discussed in questions of the extent and nature of scientific psychology, namely the distinction between the *personal* and *subpersonal* level. This distinction was originally introduced by Dennett (cf. Dennett 1969) to denote, on the one hand, discourse that concerns persons as such and their different attributes and activities: genuine subjects of thought who undertake decisions and actions, who have beliefs, desires, pains and emotions, etc.; and, on the other, discourse concerning, roughly, the enabling conditions for the personal level events and properties, which might include anything from details of neurophysiology to (seemingly) intentional descriptions of subsystems of the mind of the kind modern cognitive science deals in. Now, exactly what this distinction amounts to, how precise it is meant to be, what its significance is and where it should be drawn are delicate questions. However, for the purposes of this essay we may put them to one side, for what I have been arguing for is merely that the kinds of explanation we give of others' actions cannot coherently be viewed as other than ordinary scientific explanations. On the face of it, this claim is simply orthogonal to questions about the divide between personal and subpersonal. On the other hand, the view I have been defending may have implications for these latter questions, but there is not space to go into that here. If my view in some way undermines the significance of the personal-subpersonal divide, then I am prepared to let the chips fall where they may, though I would point out that, on my view, one is not obliged to renounce the idea that *giving* FP-explanations is personal in the sense that it may involve the first-personal use of the imagination (cf. the discussions of simulation theory in sections 2 and 3).

In view of what I have argued, the idea that there is something special about what makes us humans special, as opposed to what makes anything else in nature special, does not seem particularly engaging. I am not supposing that this conclusion will be palatable for many, or that it is in itself satisfactory as a conclusion (though neither am I denying this).

Nevertheless, that we are specially special cannot be substantiated by the idea that we are amenable to some different kind of explanation than those we find in science.<sup>23</sup>

### References

- Baron-Cohen, S. 1995 *Mindblindness*. Cambridge, Mass.: MIT Press.
- Botterill, G. & Carruthers, P. 1999 *The Philosophy of Psychology*. Cambridge: Cambridge University Press.
- Chater, N. & Pickering, M. 1997 'Two projects for understanding the mind: a response to Morris and Richardson' *Minds and Machines*, 7, 553-569.
- Child, W. 1994 *Causality, Interpretation, and the Mind*. Oxford: Oxford University Press.
- Chomsky, N. 1980 *Rules and Representations*. Oxford: Blackwell.
- \_\_\_\_\_ 1988 *Language and Problems of Knowledge*. Cambridge, Mass.: MIT Press.
- Churchland, P. 1981 'Eliminative materialism and the propositional attitudes' *Journal of Philosophy*, 78, 67-90.
- Clark, A. 1993 *Associative Engines*. Cambridge, Mass.: MIT Press.
- Crane, T. 1998 'How to define your (mental) terms' *Inquiry*, 41, 341-54.
- Davidson, D. 1980 *Essays on Actions and Events*. Oxford: Oxford University Press.
- \_\_\_\_\_ 'Problems in the explanation of action' in P. Pettit, R. Sylvan & J. Norman, eds., *Metaphysics and Morality: Essays in Honour of J.J.C Smart*. Oxford: Blackwell.
- Davies, M. 1994 'The mental simulation debate' *Proceedings of the British Academy*, 83, 99-127.
- Dennett, D. 1969 *Content and Consciousness*. London: Routledge.

---

<sup>23</sup> This paper has improved substantially since I received comments and advice on an earlier version from Petri Ylikoski of the University of Helsinki, to whom I am indebted. Its origin is notes for a graduate seminar held at the University of Oslo in the Spring of 1999, and parts of it have been presented at a seminar on philosophy of science at the University of Oslo, Autumn 2000, and at an institute seminar at NTNU in January 2001. I would like to thank the participants of all these seminars for feedback and discussion. Thanks finally to two anonymous referees for *Erkenntnis* for comments, corrections and suggestions.

*Is Folk Psychology Different?*

- Evans, G. 1973 'The causal theory of reference' reprinted in his *Collected Papers*.  
Oxford: OUP, 1982.
- Evnine, S. 1991 *Donald Davidson*. Oxford: Polity
- Fodor, J. 1974 'Special sciences' reprinted in his *Representations*. Cambridge, Mass.: MIT  
Press, 1981.
- \_\_\_\_\_ 1987 *Psychosemantics*. Cambridge, Mass.: MIT Press.
- \_\_\_\_\_ 1983 *The Modularity of Mind*. Cambridge, Mass: MIT Press.
- \_\_\_\_\_ 1998 *In Critical Condition*. Cambridge, Mass: MIT Press.
- Føllesdal, D. 1985 'Causation and explanation: a problem in Davidson's theory of action and  
mind' in E. Lepore & B. McLaughlin, eds., *Actions and Events: Perspectives on the  
Philosophy of Donald Davidson*. Oxford: Blackwell.
- Gordon, R. 1986 'Folk psychology as simulation' *Mind and Language*, 1.
- Hempel, C. 1965 'Reasons and covering laws in historical explanation' in S. Hook, ed.,  
*Philosophy and History*. New York: New York University Press.
- Henderson, D. 1993 *Interpretation and Explanation in the Human Sciences*. New York: State  
University of New York Press.
- Knowles, J. 2001 'Does intentional psychology need vindicating by cognitive science?'  
*Minds and Machines*, 11, 347-377.
- Lewis, D. 1970 'How to define theoretical terms' reprinted in his *Philosophical Papers*,  
Volume 1. Oxford: Oxford University Press, 1983.
- \_\_\_\_\_ 1972 'Psychophysical and theoretical identifications' reprinted in N. Block, ed.,  
*Readings in the Philosophy of Psychology*, Vol. 1. Cambridge, Mass.: Harvard  
University Press.
- McDowell, J. 1980 'Physicalism and primitive denotation: Field on Tarski' in M. Platts, ed.,  
*Reference, Truth and Reality*. London: Routledge.
- \_\_\_\_\_ 1985 'Functionalism and anomalous monism' in B. McLaughlin and E. Lepore, eds.,  
*Actions and Events*. Blackwell.
- \_\_\_\_\_ 1994 *Mind and World*. Cambridge, Mass.: Harvard University Press.

*Is Folk Psychology Different?*

- McGinn, C. 1979 'Problems in the explanation of action' in N. Bolton, ed., *Philosophical Problems in Psychology*. London: Methuen.
- Morris, W. & Richardson, R. 1995 'How not to demarcate cognitive science and folk psychology: A response to Pickering and Chater' *Minds and Machines*, 5, 339-335.
- Peacocke, C. 1991 *A Theory of Concepts*. Cambridge, Mass: MIT Press.
- Pickering, M. & Chater, N. 1995 'Why cognitive science is not formalised folk psychology' *Minds and Machines*, 5, 309-337.
- Pinker, S. 1997 *How the Mind Works*. Harmondsworth: Penguin.
- Putnam, H. 1975 'The meaning of "meaning"' in his *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Quine, W. 1960 *Word and Object*. Cambridge, Mass.: MIT Press
- Ruben, D. 1990 *Explaining Explanation*. London: Routledge.
- Spelke, E.: 1995 'Initial knowledge: six suggestions' *Cognition*, 50, 433-47.
- Stone, T. & Davies, M. 1996 'The mental simulation debate: a progress report' in P. Carruthers & P. Smith *Theories of Theories of Mind*. Cambridge: Cambridge University Press.