



OXFORD JOURNALS  
OXFORD UNIVERSITY PRESS

## Mind Association

---

The Humean Theory of Motivation

Author(s): Michael Smith

Source: *Mind*, New Series, Vol. 96, No. 381 (Jan., 1987), pp. 36-61

Published by: [Oxford University Press](#) on behalf of the [Mind Association](#)

Stable URL: <http://www.jstor.org/stable/2253760>

Accessed: 05/09/2011 20:36

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at  
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Oxford University Press and Mind Association are collaborating with JSTOR to digitize, preserve and extend access to *Mind*.

<http://www.jstor.org>

# *The Humean Theory of Motivation*

MICHAEL SMITH

## 1. *Two principles*

It has recently been argued that the Humean Theory of Motivation is a dogma in philosophical psychology, that the dogma is fundamentally incorrect, and that the Humean theory should therefore be replaced in philosophical psychology with a more plausible theory of motivation. I am thinking in particular of recent work by Tom Nagel, John McDowell, and Mark Platts.<sup>1</sup>

In fact the Humean seems committed to two claims about motivating reasons, a weaker and a stronger. However, there is no agreement amongst non-Humeans as to whether the weaker and the stronger are both equally unacceptable, or whether it is only the stronger that we have reason to reject. The stronger—the claim that is, as I understand it, constitutive of the Humean theory—is the claim that motivation has its *source* in the presence of a relevant desire and means-end belief. This claim finds more formal expression in the following principle:

- P1. R at t constitutes a motivating reason of agent A to  $\phi$  iff there is some  $\psi$  such that R at t consists of a desire of A to  $\psi$  and a belief that were he to  $\phi$  he would  $\psi$ .<sup>2</sup>

Non-Humeans are united in their rejection of P1. However, P1 entails the following weaker principle:

- P2. Agent A at t has a motivating reason to  $\phi$  only if there is some  $\psi$  such that, at t, A desires to  $\psi$  and believes that were he to  $\phi$  he would  $\psi$ .

—the principle that motivation requires the *presence* of a relevant desire and means-end belief—and non-Humeans are not at all united in their rejection of P2. Thus, for instance, while Tom Nagel, and John McDowell following him, have argued that P2 is acceptable because consistent with the claim that

<sup>1</sup> See Nagel, *The Possibility of Altruism*, Princeton University Press, 1970, Part 2; McDowell, 'Are Moral Requirements Hypothetical Imperatives?' in *Proceedings of the Aristotelian Society Supplementary Volume*, 1978, 'Virtue and Reason' in *The Monist*, 1979, and 'Non-cognitivism and Rule Following' in *Wittgenstein: To Follow a Rule*, London, Routledge and Kegan Paul, 1981, edited by Holtzman and Leich; Platts, *Ways of Meaning*, London, Routledge and Kegan Paul, 1979, Chapter 10, and 'Moral Reality and the End of Desire' in *Reference, Truth and Reality*, London, Routledge and Kegan Paul, 1980, edited by Mark Platts. I shall hereafter refer to the Humean Theory of Motivation as the 'Humean theory'.

<sup>2</sup> Compare Davidson's 'Actions, Reasons and Causes' in his *Essays on Actions and Events*, Oxford, Clarendon Press, 1980, p. 5.

the desires and means-end beliefs that must be present whenever there is motivation are not themselves the *source* of such motivation<sup>3</sup>—other non-Humeans, such as Mark Platts, have argued that P<sub>2</sub> is also unacceptable because either ‘phenomenologically false . . . or utterly vacuous’.<sup>4</sup>

I am inclined to agree with the non-Humeans that the Humean theory is a dogma in philosophical psychology, a ‘dogma’ in the sense that both P<sub>1</sub> and P<sub>2</sub> seem to find a fair degree of uncritical acceptance. However, unlike the non-Humeans, I do not believe that the Humean theory, as characterized by P<sub>1</sub>, is fundamentally incorrect (and thus I do not think that P<sub>2</sub> is either phenomenologically false or utterly vacuous). My task in the present paper is thus to offer an explicit argument for the Humean theory, and to defend it against the objections offered by the likes of Nagel, McDowell, and Platts. If the argument offered here is correct, then the Humean theory is the expression of a simple but important truth about the nature of motivating reasons, a truth that non-Humeans have failed to appreciate either because they have failed to distinguish motivating reasons from other sorts of reasons, or because they have an inadequate conception of desire, or because they have overlooked the implications of the fact that reason explanations are teleological.

## 2. *Motivating reasons and normative reasons*

P<sub>1</sub> is a principle connecting *motivating* reasons with the presence of desires and beliefs. We must begin by emphasizing this fact, otherwise it will seem simply implausible to suppose that P<sub>1</sub> provides individually necessary or jointly sufficient conditions for a state’s constituting a motivating reason.

In order to see this, consider the following counterexamples to the claim that P<sub>1</sub> provides necessary conditions:

- (i) Suppose I now desire to purchase an original Picasso, but I do not now believe that were I to purchase the painting before me I would do so—suppose I don’t believe that it is a Picasso. Surely it would be appropriate for an outsider to say that I have a reason to purchase the painting before me. But I lack the relevant belief.
- (ii) Suppose that I am standing on someone’s foot so causing him pain, and that I know that this is what I am doing. Surely we can imagine its being appropriate for an outsider to say that I have a reason to get off his foot even though I lacked the relevant desire, and, indeed, even if I desired to cause him pain.

Consider now the following counterexample to the claim that P<sub>1</sub> provides a sufficient condition:

<sup>3</sup> *The Possibility of Altruism*, p. 29; ‘Are Moral Requirements Hypothetical Imperatives?’, p. 15.

<sup>4</sup> *Ways of Meaning*, p. 256.

(iii) Suppose I now desire to drink a gin and tonic and believe that I can do so by mixing the stuff before me with tonic and drinking it.<sup>5</sup> Suppose further that this belief is false—the stuff before me is not gin, it is petrol. Surely it would be appropriate for an outsider to say that I had no reason to mix this stuff with tonic and drink it. Yet I have both the relevant belief and desire.

Do we have, in examples the like of these, the makings of an objection to P<sub>1</sub>, and hence to the Humean theory? We do not. The reason why was perhaps evident from the start. The outsider's perspective is not irrelevant to the examples.

It has been noticed before that the claim that A has a reason to  $\phi$  is ambiguous. It may be a claim about a *motivating* reason that A has or a claim about a *normative* reason that A has.<sup>6</sup> The crucial feature these reasons have in common is that each purports to justify certain behaviour on A's behalf; for there is an *a priori* connection between citing an agent's reasons for acting in a certain way and giving a partial justification for his acting in that way, that is, a specification of what was to be said for acting in the way in question. This is not to say that the existence of a reason for acting in a certain way, be it motivating or normative, entails that, all things considered, acting in that way is justified. But it is to say that, abstracting away from other considerations, the action is justified from the perspective of the reason (more on this below). However, in virtue of their differences, motivating and normative reasons forge the connection between justification and action differently.

The distinctive feature of a motivating reason to  $\phi$  is that in virtue of having such a reason an agent is in a state that is *potentially explanatory* of his  $\phi$ -ing.<sup>7</sup> (Note the 'potentially'. An agent may therefore have a motivating reason to  $\phi$  without that reason's being overriding.) It is thus natural to suppose that an agent's motivating reasons are, as we might put it, *psychologically real*, for it would seem to be part of our concept of what it is for an agent's reasons to have the potential to explain his behaviour that his having those reasons is a fact about *him*; that is, that the goals that such reasons embody are *his* goals.<sup>8</sup> And it is also natural, therefore, to assign to an agent's motivating reasons the minimal justificatory role possible: the role of

<sup>5</sup> The example comes from Williams, 'Internal and External Reasons' in his *Moral Luck*, Cambridge University Press, 1981, p. 102.

<sup>6</sup> I borrow these terms from *The Possibility of Altruism*, p. 4 and p. 18. In his 'Reasons for Action and Desire' in *Proceedings of the Aristotelian Society Supplementary Volume*, 1972, Michael Woods makes room for a somewhat similar distinction when he notes that 'the concept of a reason for an action stands at the point of intersection, so to speak, between the theory of the explanation of actions and the theory of their justification' (p. 189).

<sup>7</sup> It will emerge in section 4 that, in the sense in which we need to think of motivating reasons as being explanatory, we need only think of them as being teleologically explanatory; we do not need to think of them as being causally explanatory.

<sup>8</sup> Compare Christopher Peacocke's objections to instrumentalism in Chapter 8 of his *Sense and Content*, Oxford, Clarendon Press, 1983.

justifying from the perspective of the value that that very reason embodies. For a motivating reason, even when it does explain an agent's behaviour, may reveal little of value in what the agent did even from his own point of view. Consider Davidson's example of the man who has always had a yen to drink a can of paint, and who ultimately yields, but not because he thinks that doing so is really worthwhile.<sup>9</sup> None the less, knowing that he has always had a yen to drink a can of paint does provide us, and him, with a partial justification for his action, albeit a justification that justifies only from a perspective that assigns value to the drinking of a can of paint, a perspective that he himself may occupy only to the extent that he has a yen to drink a can of paint, and that we none of us may actually share.

However, to say that someone has a normative reason to  $\phi$  is to say something different. It is to say that there is some normative requirement that he  $\phi$ 's. It is therefore to justify his  $\phi$ -ing from the perspective of the normative system that generates that requirement. For present purposes there is no need to be precise about the kinds of normative requirement, and hence the kinds of normative reasons, that there may be. For all that has been said here there may therefore be as many kinds of normative reason as there are normative systems for generating reasons: normative reasons of rationality, of prudence, of morality, and perhaps normative reasons of other kinds. Nor is there any need to be precise about the relation between the normative reasons an agent has, at a time, and the motivating reasons he has at that time. Rather, the important point to note is that, on any plausible conception of what it is for there to be a normative requirement and the relation that therefore exists between this normative requirement and an agent's motivating reasons, he may well be motivated to do what he is required to do (that is, he may have a motivating reason to do what he has a normative reason to do), he may be motivated to do something that there is no normative requirement for him to do (that is, he may have a motivating reason to do what he has no normative reason to do), and there may be a normative requirement that he do what he has no motivation to do (that is, he may have a normative reason to do what he has no motivating reason to do).<sup>10</sup> Given that motivating and normative reasons may come apart, we must therefore emphasise that P1 purports to give necessary and sufficient conditions for the existence of *motivating* reasons. P1 is silent concerning the conditions under which an agent has some *normative* reason.

Consider now the examples. I said that the outsider's perspective is not irrelevant. The reason is that the outsider's perspective draws our attention

<sup>9</sup> 'Actions, Reasons and Causes', p. 4.

<sup>10</sup> Here I assume that the mere existence of the materials with which to construct a partial justification for acting in a certain way from the agent's point of view—that is, the mere existence of a motivating reason to act in that way—does not suffice for the existence of a normative reason to act in that way. If that assumption is wrong, then there will be at least one kind of normative reason for which these claims are false. That will not affect the argument that follows. For the argument requires only that there are some kinds of normative reasons for which these claims are true.

to the normative requirement in each case. Thus, note that in (i) the reason that I have to buy the painting in front of me is a normative reason. For it suffices for the truth of the claim that I have such a reason, that there is a requirement—in this case, in the broad sense, a requirement of rationality<sup>11</sup>—that I buy the painting in front of me. For I want to buy a Picasso and the painting in front of me is a Picasso. But the existence such a normative reason does not suffice for my having a motivating reason to buy the painting in front of me. For, since I do not believe that that painting is a Picasso, I am not in a state that is potentially explanatory of my buying it. (I am, of course, in a psychological state that is potentially explanatory of my buying a Picasso, for I desire to buy a Picasso and believe that were I to buy a Picasso I would buy a Picasso. But the Humean will say that this is not to have a motivating reason to buy the painting in front of me. It is rather to have a motivating reason to buy a Picasso. He will thus regard an agent's desire to  $\phi$  together with the trivial belief that were he to  $\phi$  he would  $\phi$  as the limiting case of having a motivating reason to  $\phi$ .) Thus the example in no way undermines the necessity of having a means-end belief for having a motivating reason.

A similar point applies in (ii), the case in which I have a reason to get off someone's foot when I am causing him pain. For it suffices for the truth of the claim that I have a reason to get off his foot that there exists a requirement—in this case moral—that I do not cause him pain, and that, in the present circumstances, in order to comply with that requirement I have to get off his foot. But, once again, the mere existence of this normative reason is consistent with the claim that I am not in a state that is potentially explanatory of my behaviour. (Indeed, note that this ought to be conceded even by those who think that moral reasons are rational requirements on action. For, as we have seen, rational requirements are in turn simply further normative reasons, and may thus exist in the absence of motivating reasons.) Thus this kind of example does not by itself show that having a desire is not a necessary condition for having a motivating reason.

Consider now (ii), the counterexample to the sufficiency of the condition. In what sense do I not have a reason to mix the stuff before me with tonic and drink it? Clearly, one thing we can say is that *prudence* would not require that I mix the stuff before me with tonic and drink it, for the stuff before me is petrol, and drinking petrol mixed with tonic would not be in my interests. However, I am in a state that is potentially explanatory of my mixing the stuff before me with tonic and drinking it, for I desire to drink a gin and

<sup>11</sup> In suggesting that the requirement is 'in the broad sense' a requirement of rationality I am following Williams in 'Internal and External Reasons', pp. 102–3 and Parfit in *Reasons and Persons*, Oxford, Clarendon Press, 1984, especially note 2a to Part I and pp. 117–20. For both Williams and Parfit think that a theory of rationality would tell us what an agent has reason to do, and both think that what an agent has reason to do from the perspective of rationality will depend on what an agent desires together with the *truth*. It will thus not depend on what he believes given that he may have false beliefs. (What I am calling a 'normative reason of rationality' seems to be what Williams calls an 'internal reason'.)

tonic and believe that the stuff before me is gin. Moreover, if that state did explain my doing so, we would certainly know what was to be said for doing so, from my point of view. Though prudence does not require my mixing the stuff before me with tonic and drinking it, and hence there is a sense in which I do not have a reason to do so, yet it seems entirely correct to suppose that I now have a motivating reason to do just this. So this is no counterexample to the claim that P<sub>1</sub> provides a sufficient condition for a state's constituting a motivating reason.

Let me emphasize what little I take myself to have shown here. In the light of the distinction between motivating and normative reasons, I have emphasized the fact that the Humean's is a theory about the nature of *motivating* reasons. His theory may yet be false. But it is not shown to be false simply by showing that P<sub>1</sub> fails to give necessary and sufficient conditions for the existence of normative reasons. Though I take this to be a fairly trivial point, it suffices to undermine one of Tom Nagel's principal arguments against the Humean theory in *The Possibility of Altruism*; an objection based largely on consideration of the conditions under which we would ordinarily say of someone that he has a reason. Nagel's objection centres on the Humean's explanation of prudential motivation.

### 3. *A preliminary objection: Nagel*

Prudential motivation is possible only if an agent's recognition of the fact that he will have a desire to  $\phi$  in the future somehow gives him a reason now to take steps to promote his  $\phi$ -ing then. The task of explaining this possibility takes on a particular form for a Humean. For, as we have seen, he holds that now having a motivating reason to  $\phi$  requires *presently* desiring to  $\phi$ . He must therefore explain how an agent's recognition that he *will* desire to  $\phi$  in the future gives rise to a *present* desire to promote his  $\phi$ -ing then. The Humean's answer is fairly predictable. He says that agents who are motivated by prudential considerations each have a quite general present desire to further their future interests.

However Nagel offers the following objection to the Humean's giving this answer:

The two features of the system to which I object are (a) that it does not allow the expectation of a future reason to provide by itself any reason for present action, and (b) that it does allow the present desire for a future object to provide by itself a reason for present action in pursuit of that object.<sup>12</sup>

Thus, as he points out, the following constitute possibilities under the Humean theory:

First, given that any desire with a future object provides a basis for reasons to do what will promote that object, it may happen that I now desire for the future

<sup>12</sup> *The Possibility of Altruism*, p. 39.

something which I shall not and do not expect to desire then, and which I believe there will then be no reason to bring about. Consequently I may have a reason now to prepare to do what I know I will have no reason to do when the time comes.

Second, suppose that I expect to be assailed by a desire in the future: then I must acknowledge that in the future I will have a *prima facie* reason to do what the desire indicates. But this reason does not obtain now, and cannot by itself apply derivatively to any presently available means to the satisfaction of the future desire. Thus in the absence of any further relevant desire in the present, I may have no reason to prepare for what I know I shall have reason to do tomorrow.<sup>13</sup>

The response that Nagel wants to elicit from us, faced by these examples, is that, in the first case, I have no reason to promote the future object despite my present desire, and that, in the second, I do have a reason to promote the object of my future desire despite my lacking a relevant present desire.

Myself I think that we do have this response and that we are right to. But I do not think that this fact counts against the Humean theory. In order to see that this is so, consider Nagel's own summary objection to the Humean theory's licensing such possibilities:

A system with consequences such as this not only fails to require the most elementary consistency in conduct over time, but in fact sharpens the possibilities of conflict by grounding an individual's plottings against his future self in the apparatus of rationality. These are formal and extremely general difficulties about the system, since they concern the relation of what is rational to what will be rational, no matter what source of reasons is operative.<sup>14</sup>

Thus if we accept Nagel's own diagnosis of our response to these examples—and I think we should—it emerges that examples like these fail even to touch the Humean. For, to take just the first (the second follows suit), Nagel's objection to the Humean's claim that an agent may have a *motivating* reason now to promote his  $\phi$ -ing in the future, despite the fact that he believes that he will have no motivating reason to  $\phi$  then, is that it would not be *irrational* to do so; that is, that he now has no reason from the perspective of rationality to do so. But this is to conflate the claim that an agent has a motivating reason to  $\phi$  with the claim that he has a normative reason from the perspective of rationality to  $\phi$ . The Humean is making only the first claim, not the second.

Moreover, if Nagel is right that it is irrational to promote  $\phi$ -ing in the future believing that one will then have no motivating reason to  $\phi$ , then the Humean can accept this on his own terms; by claiming that a theory of rationality requires that agents have the desire to promote their future interests.<sup>15</sup> For, importantly, the tasks of constructing a theory of motivating reasons and a theory of the normative reasons of rationality are

<sup>13</sup> *The Possibility of Altruism*, pp. 39–40.

<sup>14</sup> *Ibid.*, pp. 40–1.

<sup>15</sup> Compare *Reasons and Persons*, pp. 131–6.



just different tasks. I therefore do not see that the rationality of prudence makes for an especial difficulty with the Humean's theory of motivation. (It may indeed provide a problem for Hume's own theory of rationality.)<sup>16</sup>

#### 4. *Why believe the Humean theory?*

We have seen that we will find no easy refutation of P1—the claim that motivating reasons are constituted by the presence of desires and means-end beliefs—by reflecting on those cases in which we would ordinarily say of someone that he has a reason to  $\phi$ . But can we find some reason actually to believe this claim?

John McDowell has attempted to diagnose commitment to the Humean theory in the following terms. He begins by isolating what he takes to be the distinctive feature of the Humean's theory; namely, that 'to cite a cognitive propositional attitude', that is, a belief, 'is to give at most a partial specification of a reason for acting; to be fully explicit, one would need to add a mention of something non-cognitive, a state of the will or a volitional event' or, in the terms in which we have put it, a desire.<sup>17</sup> He then goes on:

I suspect that one reason people find . . . [this claim] obvious lies in their inexplicit adherence to a quasi-hydraulic conception of how reason explanations account for action. The will is pictured as the source of forces that issue in the behaviour such explanations explain. This idea seems to me a radical misconception of the sort of explanation a reason explanation is, but it is not my present concern.<sup>18</sup>

I am not sure that I understand McDowell's diagnosis here. But, in so far as I do, it seems to me to get things entirely wrong.<sup>19</sup>

According to McDowell, one reason people believe the Humean theory is that they have a 'quasi-hydraulic' conception of how reason explanations account for action; that is, in less prolix terms, because they have a *causal* conception of reason explanations. Moreover, in McDowell's view, this lays the Humean theory so supported open to an objection; for, he says, a causal conception of reason explanations is a 'radical misconception'. Let me begin with the second point first.

Those of us who do not agree that causal conceptions are radically misconceived may well think that McDowell has here provided *us* with good reasons for believing the Humean theory; for McDowell suggests that the Humean theory is supported by something that we believe to be true. However, it seems to me that we would be conceding too much to McDowell if we were to argue in this way. For, now taking up his first point, I doubt that there is any support to be found for the Humean theory in a causal conception of reason explanations.

<sup>16</sup> This seems to be Nagel's real objection to the Humean on pp. 64–5 of *The Possibility of Altruism*.

<sup>17</sup> 'Non-cognitivism and Rule Following', p. 154.

<sup>18</sup> *Ibid.*, p. 155.

<sup>19</sup> Though, in fairness to McDowell, see my comments on Hume's own view in section 6, and footnote 38 below.

In order to see this we only need ask why a causal conception should be thought to support especially the Humean theory. To be sure, one who holds that reason explanations are causal must conceive of some psychological states as possessed of *causal force*. But why, as McDowell seems to assume, must he think that *desires* are the only psychological state possessed of causal force? Why mightn't he think instead that only certain *beliefs* are possessed of causal force? McDowell offers no argument on this point.

Indeed, when we consider the reason causal theorists actually give for holding a causal conception, it emerges that no such argument is forthcoming. For they reason roughly as follows: 'We ordinarily say of agents that they  $\phi$  *because* they have reason to  $\phi$ . The "because" here may uncontroversially be regarded as the "because" of rationalization; or, better, the "because" of teleological explanation. But now observe that an agent may have reason to  $\phi$  and  $\phi$ , and yet not  $\phi$  because he has reason to  $\phi$ . What then is the feature that makes the difference between this case and the case in which the agent  $\phi$ s *because* he has reason to  $\phi$ ? The only illuminating answer available is that the reasons in the second case *cause* the agent to  $\phi$ .'<sup>20</sup> It thus emerges that the argument causal theorists give for a causal conception of reason explanations makes no substantial assumption about the nature of the reasons we have. So, it seems, we should be able to accept or reject this argument quite independently of our views concerning the nature of reasons. The upshot is that if Humeans and non-Humeans alike may have a causal conception of reason explanations then it cannot be that holding a causal conception supports especially the Humean theory.

I think this shows that the Humean is engaged in a debate that is both independent of and more fundamental than the debate over whether reason explanations are causal. In short the difference is this. The causal and non-causal theorist can both accept that reason explanations are teleological explanations without enquiring further into what it is about the nature of reasons that makes it possible for reason explanations to be teleological explanations—that is, explanations that explain by making what they explain intelligible in terms of the pursuit of a goal. For their disagreement concerns the further question whether such explanations are themselves in turn a species of causal explanation; a disagreement which may, as I have suggested, cut across disagreements concerning the nature of reasons. But, as I see it, the Humean and non-Humean are precisely engaged in a dispute concerning what it is about the nature of reasons that makes it possible for reason explanations to be teleological explanations. If this is right, then it would seem that there will be only one reason to believe the Humean's theory, if indeed we should believe his theory at all, and that is that the Humean's theory is alone able to make sense of motivation as the pursuit of a goal.

I want to argue that this is indeed the case in the remainder of this paper.

<sup>20</sup> This is a summary of the argument in 'Actions, Reasons and Causes', pp. 8–11, the *locus classicus* of arguments for a causal conception of reason explanations.

My reason for believing this is relatively simple—it seems to me to follow from a proper conception of desire. I therefore proceed by focussing on two different conceptions of desire. One of these gives no support to the Humean theory. I argue that this conception is anyway implausible. Perhaps unsurprisingly, this seems to be the conception of desire held by many opponents of the Humean theory. There is, however, an alternative and more plausible conception. This conception enables us to see that desires must be constituents of reasons given that reasons must themselves be constituted by goals.

### 5. *Desires and phenomenology*

According to Hume, desires are a species of the passions, and passions are, in turn, a certain kind of *feeling*. Hume seems to hold that this is so not just in the trivial sense that passions are a species of perception and perceptions are a kind of feeling.<sup>21</sup> Rather he seems to be suggesting that when we desire something ‘we feel an . . . emotion of aversion or propensity’;<sup>22</sup> as though, as Stroud puts it when discussing Hume’s conception of desire, his view is that we are ‘*directly aware*’ of the presence of the desires that we have.<sup>23</sup>

Hume’s suggestion is not entirely misguided. For there is such a thing as the phenomenology of desire; as, for instance, to use one of Hume’s own examples, ‘when I am angry I am possest with the passion’.<sup>24</sup> That is, we may agree with Hume that, on occasion, when I have a desire, I am possessed with a psychological feeling; an analogue of a bodily sensation. This may suggest an elaboration of Hume’s view. For if we take quite seriously his suggestion that all desires are known by the way they make us feel, then, in an attempt to explain why this is so, we may be led to identify desires with such psychological feelings. And this may in turn lead us to endorse what I shall call the ‘strong phenomenological conception’ of desires; the view that desires are, like sensations, simply and essentially states that have a certain phenomenological content.

Perhaps unsurprisingly, I think that the strong phenomenological conception of desires ought to be rejected. For it seems to me that there is no way such a conception can be married with a plausible epistemology of desire.<sup>25</sup>

Now I suspect that there will be some who think that this objection doesn’t even get off the ground. For they will say: ‘Surely the strong

<sup>21</sup> See *Treatise of Human Nature*, Oxford, Clarendon Press, 1958, p. 190.

<sup>22</sup> *Treatise*, p. 414.

<sup>23</sup> See Barry Stroud, *Hume*, London, Routledge & Kegan Paul, 1977, p. 163.

<sup>24</sup> *Treatise*, p. 415.

<sup>25</sup> I suspect that there will be some who think that this is all too obvious. They will think that the strong phenomenological conception is a strawman and thus not worth considering. However, I disagree. I think that it is worth working through our objections to strong phenomenological conceptions in order better to understand our objections to phenomenological conceptions quite generally. Those who disagree may prefer to skip the next six paragraphs.

phenomenological conception of desire makes the epistemology of desire unproblematic. For the epistemology of desire becomes like the epistemology of sensation. Thus, just as it is plausible to hold that a subject is in pain if and only if he believes that he is in pain—for we take it that a subject is in a state with a certain phenomenological content if and only if he believes himself to be in a state with that content—so, if we think of desires on the model of sensations, it is plausible to hold that a subject desires to  $\phi$  if and only if he believes that he desires to  $\phi$ .<sup>26</sup> What exactly is wrong with this? There are two things wrong with it. I begin with an objection that concedes more to the strong phenomenological conception than strictly ought to be conceded. Doing so teaches us a valuable lesson.

As I understand it, the principle that a subject desires to  $\phi$  if and only if he believes that he desires to  $\phi$  is supposed to express a necessary truth; the putative truth that we are infallible about what we desire. But, intuitively at any rate, this principle is simply false. Thus, conceding for the moment that the strong phenomenological conception does entail that a subject desires to  $\phi$  if and only if he believes that he desires to  $\phi$ , it ought to be rejected. I argue by counterexample.

Suppose each day on his way to work John buys a newspaper at a certain newspaper stand. However, he has to go out of his way to do so, and for no apparently good reason. The newspaper he buys is on sale at other newspaper stands on his direct route to work; there is no difference in the price or condition of the newspapers bought at the two stands; and so on. There is, however, this difference between the stands. There are mirrors behind the counter of the stand where John buys his newspaper. Given their placement, one who buys a newspaper there cannot help but look at himself. Let's suppose, however, that if it were suggested to John that the reason he buys his newspaper at that stand is that he wants to look at his own reflection, he would vehemently deny it. And it wouldn't seem to John as if he were concealing anything in doing so. However, finally, let's suppose that if the mirrors were removed from the stand, his preference for that stand would disappear. If all this were the case, wouldn't it be plausible to suppose that John in fact desires to buy his newspaper at a stand where he can look at his own reflection; that, perhaps, he has a narcissistic tendency and that buying his newspaper at that stand enables him to indulge it on the way to work? And wouldn't it also be plausible to suppose that he does not believe that this is so, given his, from his point of view, sincere denials? If this is agreed, then we have reason to reject the principle left to right.<sup>26</sup>

Consider another example. Suppose John professes that one of his fundamental desires is to be a great musician. However, his mother has

<sup>26</sup> Indeed, it seems to me that there are more mundane counterexamples to the principle left to right. Consider cases in which you go to the refrigerator convinced that there is something in particular that you want, though you aren't quite sure what it is. Then, while looking at the contents you suddenly, as we should put it, 'realise what it was that you wanted all along'. If we wish to respect this common sense description of such occurrences then we should reject the principle left to right.

always drummed into him the value of music. She is a fanatic with great hopes for her son's career as a musician; hopes so great that she would be extremely disappointed if he were even less than an excellent musician, let alone if he were to give up music altogether. Moreover, John admits that he has a very great desire not to upset her, though he denies that this in any way explains his efforts at pursuing excellence in music. However, suppose now that John's mother dies, and that, upon her death, he finds that all of his interest in music vanishes. He gives up his career as a musician and pursues a career in film as an actor. In such circumstances, wouldn't it be plausible to suppose that John was just mistaken about what he originally wanted to do and that, despite the fact that he believed that achieving excellence in music was a fundamental desire of his, it never was? If so, then we have reason to believe that the principle is false right to left as well.

If this is agreed, then any conception of desires that entails that a subject desires to  $\phi$  if and only if he believes that he desires to  $\phi$  is a conception that is to be rejected, rejected for the simple reason that the epistemology it provides is implausible. Thus, if the phenomenological conception entails such a principle, it ought to be rejected. This teaches us the following lesson. It is an adequacy constraint on any conception of desire that the epistemology of desire it recommends allows that subjects may be fallible about the desires they have.

I said that the first objection concedes more to the strong phenomenological conception than strictly ought to be conceded. What it concedes is that the strong phenomenological conception does entail that a subject desires to  $\phi$  if and only if he believes that he desires to  $\phi$ . But this it most certainly does not do. Here, then, is a second objection.

Let's grant for the moment that desires are like sensations in that they essentially have phenomenological content. It must be noted that they differ from sensations in that they have in addition *propositional* content. For ascriptions of desires, unlike ascriptions of sensations, may be given in the form 'A desires that p', where 'p' is a sentence. Thus, whereas A's desire to  $\phi$  may be ascribed to A in the form *A desires that he  $\phi$ s*, A's pain cannot be ascribed to A in the form *A pains that p*.<sup>27</sup>

It is therefore ambiguous to claim that the epistemology of desire is 'like the epistemology of sensation'. To be sure, if desires are essentially phenomenological states, then the epistemology of the phenomenological content of a desire may be based on the epistemology of sensation. But what about the epistemology of the propositional content of desire? This cannot be based on the epistemology of sensation at all, for sensations have no propositional content. It therefore turns out that we have an even stronger reason to reject the strong phenomenological conception of desire. For, according to this conception, there is simply no difference between desires and sensations. Each is a state which simply and essentially has

<sup>27</sup> Compare 'Moral Reality and the End of Desire', pp. 74-7.

phenomenological content. The strong phenomenological conception of desires is thus unable to account for the fact that desires *have* propositional content *at all*. Little wonder that it cannot provide a plausible epistemology of the propositional content of desires.

I suspect that, for this very reason, some will have thought that the strong phenomenological conception of desire was a straw man all along. But note that, with our objections to the strong phenomenological conception firmly in place, we are now in a position to argue against all versions of the phenomenological conception, even the more plausible weaker conceptions according to which desires are *like* sensations in that they have phenomenological content essentially, but *differ* from sensations in that they have propositional content as well. For we can now say this about all such conceptions: they in no way contribute to our understanding of what a desire as a state with propositional content is, for they cannot explain how it is that desires have propositional content; they therefore in no way explain the epistemology of the propositional content of desire; and they thus require supplementing by some independent and self-standing account of what a desire is which explains how it is that desires have propositional content and which explains how it is that we have fallible knowledge of what it is that we desire.

The question that immediately arises with regard to weaker phenomenological conceptions is then why we should believe that any such conception is true. The only answer available is that a phenomenological conception is alone true to the phenomenology of desire. But is this answer plausible? Do we really believe that desires are states that have phenomenological content essentially? That is, do we believe that if there is nothing that it is like to have a desire, at a time, then it is not being had at that time?

I should say that, at least as far as common sense opinion goes—and what else do we have to go by in formulating a philosophical conception of folk psychological states?—we evidently have no such belief. Consider, for instance, what we should ordinarily think of as a long term desire; say, a father's desire that his children do well. A father may actually feel the prick of this desire from time to time; in moments of reflection on their vulnerability, say. But such occasions are not the norm. Yet we certainly wouldn't ordinarily think that he loses this desire during those periods when he lacks such feelings. Or consider more mundane cases like those that Stroud mentions in his discussion of Hume, cases in which, as we should ordinarily say, I desire to cross the road and do so, or in which I desire to write something down and so write it down. As Stroud points out, in such cases 'it is difficult to believe that I am overcome with emotion . . . I am certainly not aware of any emotion or passion impelling me to act'; rather 'they seem the very model of cool, dispassionate action'.<sup>28</sup> However, it would be grossly counter to our common sense opinion to conclude that simply because I do

<sup>28</sup> Hume, p. 163.

not introspect the presence of desires in such cases so I incorrectly attribute desires to myself—that I cross the road and write things down even though I do not want to!

Of course, if we thought that there was nothing for a desire to be, in the absence of its being felt, then we might, in our role of philosophical theorist, feel ourselves forced into concluding that some of our common sense attributions are mistaken, and hence feel ourselves forced into revising our common sense opinions in favour of a phenomenological conception. But given that a phenomenological conception is unable to deliver an account of desire as a state with propositional content, we should feel no such pressure in our role of philosophical theorist. Rather we should concede that a desire may be had in the absence of its being felt.

This is significant. For many non-Humeans seem to work with a phenomenological conception of desire, and then use the fact that we do not introspect the presence of desires whenever there is motivation against the Humean theory.

Consider, for instance, the following argument of Mark Platts's:

The crucial premiss . . . is the claim that any full specification of a reason for an action, if it is to be a reason for the potential agent for action, must make reference to that agent's desires. At first sight, it seems a painful feature of the moral life that this premiss is false. We perform many intentional actions in that life that we apparently do not desire to perform. A better description of such cases appears to be that we perform them because we think them desirable. The difficulty of much of moral life then emerges as a consequence of the apparent fact that desiring something and thinking it desirable are both distinct and independent.

The premiss can, of course, be held true by simply claiming that, when acting because we think something desirable, we do indeed desire it. But this is either phenomenologically false, there being nothing in our inner life corresponding to the posited desire, or utterly vacuous, neither content nor motivation being given to the positing of the desire. Nothing but muddle (and boredom) comes from treating desire as a mental catch-all.<sup>29</sup>

Thus, according to Platts, the Humean may hold that when we believe that something is desirable we do desire it. However, if he does, then Platts claims that he is impaled on the horns of a dilemma. But consider the horns of Platts's dilemma.

If there is no reason why *any* theorist should accept a phenomenological conception of desire, as we have seen that there is not, then it can hardly be an objection to the Humean's theory that we are unable to introspect the presence of each and every desire that he says we have. Thus we should not force the Humean onto the phenomenological falsehood of the first horn. Platts might agree. But, he would say, this merely forces the Humean onto the second horn of his dilemma. Here Platts seems to claim that if we do not accept a phenomenological conception of desire, then the positing of a desire

<sup>29</sup> *Ways of Meaning*, p. 256.

must be ‘utterly vacuous’, or without ‘content’.<sup>30</sup> But, as we have seen, given just the assumption that desires are states with propositional content, an assumption that must be accepted even by one who does wish to endorse a phenomenological conception, the *only* way that we can give content to the concept of such a state, and hence to the positing of a desire, is precisely via some independent and self-standing non-phenomenological conception. So, far from non-phenomenological conceptions making ascriptions of desire ‘utterly vacuous’, non-phenomenological conceptions alone make the ascription of desires with propositional contents possible.

Indeed, even John McDowell, who himself rejects a phenomenological conception, covertly ascribes such a conception to the Humean when arguing against him.<sup>31</sup> This emerges in McDowell’s defence of his own view that the virtuous agent may be motivated by his conception of the situation in which he finds himself, something that, according to McDowell, may properly be thought of as a cognitive state. McDowell rightly supposes that the Humean would respond that, if someone who has such a conception is indeed motivated, then getting him to have such a conception must involve getting him to have a certain desire. But he then interprets this as the suggestion that

‘See it like this’ is really a covert invitation to feel, quite over and above one’s view of the facts, a desire which will combine with one’s belief to recommend acting in the appropriate way.<sup>32</sup>

And he rightly rejects this suggestion. But, unless it is compulsory to accept a phenomenological conception of desire, why give the Humean’s response that interpretation? To be sure, getting someone to have a certain view of the facts may not involve getting him to *feel* a certain desire, but it may involve getting him to *have* a certain desire none the less.<sup>33</sup>

## 6. *Desires, directions of fit, and dispositions*

We have seen that there must be an alternative to phenomenological conceptions of desire, an alternative that allows us to make sense of desires as states with propositional contents and that thus allows us to make sense of our common sense desire attributions. But what is the alternative to be?

<sup>30</sup> I say ‘seems’, for Platts does in fact consider an alternative characterization of desires. I discuss the alternative characterization in section 6. I discuss Platts’s assessment of the support that the alternative characterization gives to the Humean theory in section 8 (ii).

<sup>31</sup> I say McDowell rejects a phenomenological conception. He nowhere says that he rejects such a conception. But given that he thinks that ‘consequentially ascribed desires are indeed desires’ (‘Are Moral Requirements Hypothetical Imperatives?’, p. 25), he must. For the point of consequentially ascribed desires is that there may be no phenomenological ground for their ascription. The idea of ‘consequentially ascribed desires’ derives from the work of Tom Nagel. For a discussion of the idea that some desires are merely consequentially ascribed see section 8 (i).

<sup>32</sup> ‘Are Moral Requirements Hypothetical Imperatives?’, p. 22.

<sup>33</sup> I briefly discuss McDowell’s view further in section 8 (iii).



Surprisingly enough, Platts himself outlines the alternative I favour; a suggestion about the difference between beliefs and desires that he attributes to Anscombe. Platts's own summary is so succinct and makes the idea sound so plausible that I shall merely quote it. (I consider below why Platts is himself subsequently so unsympathetic towards the idea.)

Miss Anscombe, in her work on intention, has drawn a broad distinction between two *kinds* of mental states, factual belief being the prime exemplar of one kind and desire a prime exemplar of the other . . . The distinction is in terms of the *direction of fit* of mental states with the world. Beliefs aim at the true, and their being true is their fitting the world; falsity is a decisive failing in a belief, and false beliefs should be discarded; beliefs should be changed to fit with the world, not vice versa. Desires aim at realisation, and their realisation is the world fitting with them; the fact that the indicative content of a desire is not realised in the world is not yet a failing *in the desire*, and not yet any reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa.<sup>34</sup>

Myself I think that this characterization of the difference between beliefs and desires captures something quite deep in our thought about their nature. Moreover, I want eventually to argue that the idea that desires are states with which the world must fit allows us to bring out an important connection between our concepts of desire and motivation.

However, as Platts notices, talk of the direction of fit of a state is highly metaphorical. This is problematic. For it seems that we would be unjustified in appealing to the concept in characterizing desires, and in illuminating the connection between desires and motivation, if we had no way of understanding it in non-metaphorical terms. Moreover, as once again Platts notices, if we take the characterization quite strictly, it is unclear whether it allows us to characterize desires at all. For, he claims, since 'all desires appear to involve elements of belief', desires are not states whose direction of fit is entirely of the second kind: the question arises whether there are any such states.<sup>35</sup>

It seems to me that Platts is right to highlight these problems with the metaphor, but that we would be wrong to think that the problems he raises are insurmountable. For I want to suggest that the metaphorical characterization of desires as states which are such that the world must fit with them meshes with another, and more plausible, suggestion about the epistemology of desires; a suggestion inspired, ironically enough, by certain other remarks of Hume's.

Hume realized all too well that alongside the 'violent passions' that affect the subject who has them, there are 'calm passions', passions that lack phenomenological content altogether. Hume was therefore cognizant of the fact that his official line on the epistemology of desires—that they are known by their phenomenology—was totally inadequate as an account of

<sup>34</sup> *Ways of Meaning*, pp. 256–7.

<sup>35</sup> *Ibid.*, p. 257.

the epistemology of the calm passions, and that he therefore needed an alternative account of the epistemology appropriate for them. As a result, Hume suggested that, by contrast with the violent passions, the calm passions 'are more known by their effects than by their immediate sensation'.<sup>36</sup>

When Stroud considers this suggestion, he points out that it commits Hume to the view that desires are to be conceived of as *the causes of actions*.<sup>37</sup> It might therefore be thought that Hume's suggestion should be of little interest to us. For I have argued that the only argument for the Humean theory, if there is to be one at all, will be that it alone is able to make sense of reason explanation as a species of teleological explanation, and that one may accept that reason explanations are teleological without accepting that reason explanations are causal (see section 4). Yet if we accept this conception of desire we immediately lock ourselves into a causal conception.<sup>38</sup>

However, though this makes acceptance of Hume's suggestion as it stands inappropriate, it seems to me that we would be wrong to abandon Hume's suggestion altogether. For if we are less interested in *Hume's* view than in a *Humean* view then it seems to me that we can find in Hume's suggestion about the epistemology appropriate for the calm passions, the inspiration for a somewhat different conception of desires; a conception that allows us to remain neutral about whether desires are causes.

According to this alternative conception, desires are states that have a certain *functional role*. That is, according to this conception, we should think of the desire to  $\phi$  as that state of a subject that grounds all sorts of his dispositions: like the disposition to  $\phi$  in conditions C, the disposition to  $\phi$  in conditions C', and so on (where, in order for conditions C and C' to obtain, the subject must have, *inter alia*, certain beliefs). For Hume's suggestion about how the calm passions are known may then be translated into the thought that the epistemology of desire is simply the epistemology of dispositional states—that is, the epistemology of such counterfactuals. This does not commit us, *as Humeans*, to the thesis that desires are to be conceived of as the causes of actions. For it is a substantial philosophical thesis to move from the claim that desires are dispositions to the claim that desires are causes.

A dispositional conception of desires enables us to solve many of the problems that we have confronted so far. For instance, a dispositional conception is precisely an account of what a desire is that explains how it can be that desires have propositional content, for the propositional content of a desire may then simply be determined by its functional role. (I say that it may 'simply' be determined by its functional role. But of course, there need be nothing simple about the functional theory that determines content.)

<sup>36</sup> *Treatise*, p. 417.

<sup>37</sup> *Hume*, p. 165.

<sup>38</sup> Perhaps this link between Hume's own conception of desire and a causal conception of reason explanations explains why McDowell thinks that the Humean theory finds support from a causal conception (see section 4).

A dispositional conception of desires also meets the constraint on the epistemology of desire argued for earlier; that the epistemology of desire allows that subjects may be fallible about the desires that they have.<sup>39</sup> For, given just the assumption that desires are dispositions to act in certain ways under certain conditions, it is implausible to suggest quite generally that if the counterfactuals that are thus true of a subject who desires to  $\phi$  are true of him then he believes that they are, and it is likewise implausible to suggest quite generally that if a subject believes that such counterfactuals are true of him, then such counterfactuals are true of him.<sup>40</sup> Furthermore, a dispositional conception of desires is consistent both with the claim that certain desires have phenomenological content essentially and with the claim that certain desires lack phenomenological content altogether. For, according to this conception, desires have phenomenological content just to the extent that the having of certain feelings is *one* of the things that they are dispositions to produce under certain conditions. Some desires may be dispositions to have certain feelings under all conditions: these have phenomenological content essentially. Other desires, though they are dispositions to behave in certain ways, may not be dispositions to have certain feelings at all: these lack phenomenological content altogether.

We are also able, given a dispositional conception of desires to see why Platts is right that, in myriad ways, desires 'involve elements of belief'. For if the desire to  $\phi$  is a certain sort of complex dispositional state of the kind described then desiring to  $\phi$  may 'involve' elements of belief in each of the following ways: the obtaining of the conditions in which the subject  $\phi$ s may require that he has certain beliefs; the truth of the counterfactual 'Were the subject in conditions C he would  $\phi$ ' may require that the subject has certain other beliefs due to holistic constraints on desire attribution; and so on. Indeed, if we take Platts's own claims about the phenomenology of desire quite seriously, then we might suggest that desires involve beliefs in other more direct ways. For, if we agree with Platts that oftentimes when we act our only awareness of the desires we have comes via our beliefs concerning the desirability of various options then we might think that this too ought to be reflected in the kind of disposition a desire is. Thus, we might say, the desire to  $\phi$  is also, *inter alia*, a disposition to believe, under certain

<sup>39</sup> Here I am grateful to Frank Jackson.

<sup>40</sup> Thus, I contend that a subject's false belief that he desires to  $\phi$  is not a state that is potentially explanatory of his behaviour. Everyday experience supports this contention. Reflect on occasions when you stand at the edge of a cold swimming pool thinking that you desire to jump in. On some such occasions your body is totally unresponsive to the desires that you profess to have. As you stand there motionless you sometimes come to the conclusion that, contrary to what you thought, you didn't really want to go swimming after all. Thus, just as we would expect if this contention were true, there are cases in which a subject believes that he desires to  $\phi$  right up until the time that he is supposed to act only to discover that he in fact has no such desire when his body fails to respond to his desire. Of course, we can construct cases in which it might *appear* that an agent's false beliefs about what he wants motivate him. The example of John the musician may perhaps be such a case. But I should claim that in such cases, as in John's, we will find that the appearance is misleading, and that there is in fact some other desire that does the motivating.

conditions, that  $\phi$ -ing is *prima facie* desirable (= 'is desired by me'); or that there is a reason to  $\phi$ ; and so on.

Finally, a dispositional conception enables us to see why, despite the fact that in these many ways desires may involve elements of belief, we may properly say of one who has a desire that he is in a state *with which the world must fit*. Moreover, and for the Humean's particular purposes perhaps more importantly, a dispositional conception of desires enables us to cash the metaphor characterizing beliefs and desires in terms of their direction of fit, and therefore draws support from it. For the difference between beliefs and desires in terms of direction of fit comes down to a difference between the counterfactual dependence of a belief and a desire that  $p$ , on a perception that *not*  $p$ : roughly, a belief that  $p$  is a state that tends to go of existence in the presence of a perception that *not*  $p$ , whereas a desire that  $p$  is a state that tends to endure, disposing the subject in that state to bring it about that  $p$ . Thus, we may say, attributions of beliefs and desires require that different *kinds* of counterfactuals are true of the subject to whom they are attributed. We may say that this is what a difference in their directions of fit *is*.

These are important results. For they serve to make a dispositional conception of desire attractive *period*, quite independent of the theory of motivation that one happens to favour. Moreover, they license us to talk unashamedly of desires as states with which the world must fit, for such talk, though metaphorical, captures the feature that distinguishes desires from beliefs; that is, such talk, though metaphorical, aptly describes the kind of dispositional state that a desire is.

The Humean's reasons for believing P1—the principle that a motivating reason is constituted by the presence of a desire and a means–end belief—may now be stated rather simply. Given that, as we have seen, all theorists should accept a dispositional conception of desires, and given that this conception licenses us to talk of desires as states with which the world must fit, the Humean's reasons are also, I think, both intuitive and compelling.

## 7. Desires, directions of fit, goals and motivating reasons

What is it for someone to have a motivating reason? The Humean replies as follows. We understand what it is for someone to have a motivating reason at a time by thinking of him as, *inter alia*, having a goal at that time; the '*alia*' here includes having a conception of the means to attain that goal. That is, having a motivating reason just *is*, *inter alia*, having a goal. But what kind of state is the having of a goal? It is a state with which *direction of fit*? Clearly, the having of a goal is a state *with which the world must fit*, rather than *vice versa*. Thus having a goal is being in a state with the direction of fit of a desire. But since all that there is to being a desire is being a state with the appropriate direction of fit, it follows that having a goal just *is* desiring. In

short, then, the Humean believes  $P_1$  because  $P_1$  is entailed by the following three premises:

- (1) Having a motivating reason *is, inter alia*, having a goal
  - (2) Having a goal *is* being in a state with which the world must fit
- and
- (3) Being in a state with which the world must fit *is* desiring.

Simple though it is, this argument is, I think, really quite powerful. After all, which premise in the argument could plausibly be denied? Let's consider them in turn.

Given just the assumption that reason explanations are teleological explanations (see section 4), (1) seems unassailable; indeed it has the status of a conceptual truth. For we understand what it is for someone to have a motivating reason in part precisely by thinking of him as having some goal.<sup>41</sup> (2) is likewise unassailable. For learning that the world is not as the content of your goal specifies is not enough for giving up that goal, but rather puts pressure on you to change the world. The most vulnerable premise is perhaps (3), the claim that being in a state with which the world must fit is desiring. I can imagine two sorts of objection to this premise. But neither objection goes very far.

First, according to Platts, Anscombe claims only that desire is 'a prime exemplar' of those states with which the world must fit. But, he might say, there are other states that have this direction of fit as well: hopes, wishes, and the like. Therefore, given (1) and (2), we should surely say that such states may constitute the having of goals as well. But, *ex hypothesi*, hopes and wishes are not desires.

However, an attack of this kind on the Humean's argument is clearly not an attack on the *spirit* of his argument, it is rather an attack on the *details* of his argument. The Humean may therefore concede the details to the objector. That is, if *desire* is not a suitably broad category of mental state to encompass all of those states with the appropriate direction of fit, then the Humean may simply define the term 'pro-attitude' to mean 'psychological state with which the world must fit', and then claim that motivating reasons are constituted, *inter alia*, by pro-attitudes.<sup>42</sup>

A second objection to (3) tackles the assumption that there are only states with one or the other direction of fit. Thus, it might be asked, why couldn't there be a state with *both* directions of fit; a state which is both such that the world must fit with it, and such that it must fit the world? If such a state were possible then, it might be said, it could constitute the having of a goal.

<sup>41</sup> See section 8 (i) for a discussion of (1) in relation to Nagel's views.

<sup>42</sup> Compare Davidson on the difference between desires and pro-attitudes in 'Actions, Reasons and Causes', p. 4.

But it would not be a desire, nor even a pro-attitude; for desires and pro-attitudes have only one direction of fit.<sup>43</sup>

However, though it may sound like a coherent possibility that there should be such a state, it really isn't. For, as we have understood the concept of direction of fit, the direction of fit of a state with the content that *p* is determined, *inter alia*, by its counterfactual dependence on a perception with the content that *not p*. A state with both directions of fit would therefore have to be such that *both*, in the presence of such a perception it tends to go out of existence, *and*, in the presence of such a perception, it tends to endure, disposing the subject that has it to bring it about that *p*. Taken quite literally, then, the idea that there may be a state with both directions of fit is just plain incoherent.

Of course, a more subtle objector might find fault with this reply. He might say that the reply works only if we take the suggestion that there is a state with both directions of fit quite literally. But, he might well ask, why take it so literally?

Indeed, he might go on the offensive. For he might suggest that our resistance to the suggestion shows the extent to which we are in the grip of what might be called an 'austere' psychological theory; a theory which, as far as the explanation of action goes, both makes use of and makes do with the concepts of belief and desire: states having the one or the other direction of fit, as that has been defined here. The objector may well admit that an austere theory can explain some of the phenomena. But he might insist that we recognize that it cannot explain all of the phenomena; that there are certain goings on that we can explain only if we enrich our psychological theory with the concept of a state that has, in a more relaxed sense, *both* directions of fit. He might cite the example of a moral 'quasi-belief' that *x* is good ('quasi-belief' because, as we shall see, it is no ordinary belief). For, he might point out, since a subject who has such a quasi-belief tends to go out of this state when presented with a perception that *x* is not good, this makes it appropriate to describe a moral quasi-belief as being such that it must fit the world. But since a subject's having the moral quasi-belief that *x* is good disposes him to promote *x*, this makes it appropriate to describe such a quasi-belief as being, in a more relaxed sense, such that the world must fit with it. Indeed, he might go on to insist that since the factor that determines the kinds of concept our psychological theories can make use of is the evidence that needs to be explained by our theories, so the example just given shows that we positively have reason to enrich our austere psychological theory with the concept of a quasi-belief: a state that is both belief-like and desire-like though identical with neither. For the evidence—our moral practices—can only be explained by the richer theory.

This more subtle objection needs careful handling. I do not know whether the richer theory is really coherent—no surprise, since to my

<sup>43</sup> David McNaughton put this objection to me.

knowledge no-one has actually attempted to formulate such a theory. But I do not think that we need to follow the objection that far down the line in order to find fault with it. For, as I see it, the problem with the objection is that it requires that we *revise* our psychological theories without proper warrant. If I have correctly described the objection, its success depends entirely on the claim that there is evidence that cannot be explained by an austere psychological theory. But the evidence the objector cites can be easily accommodated by the resources of an austere theory; that is, within the accepted framework of beliefs and desires.

In order to see that this is so, consider the following parody of the argument just given: 'There is evidence for the existence of state that is belief-like and desire-like but identical with neither. Consider, for instance, the "quasi-belief" that  $\phi$ -ing is *prima facie* desirable ("quasi-belief" because, as we shall see, it is no ordinary belief). This state is appropriately described as having both directions of fit because, first, a subject tends to go out of this state if he is presented with a perception with the content that  $\phi$ -ing is not *prima facie* desirable, and, second, because a subject's being in such a state disposes him to  $\phi$ .' But this argument is hopeless. For, within the confines of our austere psychological theory, we may say that there are beliefs that a subject can have only if he has certain desires. Thus, we may say, a subject's quasi-belief that  $\phi$ -ing is *prima facie* desirable is best thought of as being such a state. It is a genuine belief because it is a state that must fit the world; that is (since 'is *prima facie* desirable' = 'is desired by me') because its truth requires that  $\phi$ -ing has the property of being desired by him. But it is also a belief the condition for having which is that a subject desires to  $\phi$  (*modulo* fallibility about one's desires). We can thus explain why one who believes that  $\phi$ -ing is *prima facie* desirable is generally disposed to  $\phi$  by noting that the normal condition for having such a belief is desiring to  $\phi$ , and desiring to  $\phi$  is, *inter alia*, a disposition to  $\phi$ .<sup>44</sup>

We therefore do not need to introduce a state of some further, mysterious,

<sup>44</sup> Does the 'normal' provide a problem here? I think not. The argument succeeds just in case it shows that there is a necessary connection of a certain kind between having the belief that  $\phi$ -ing is *prima facie* desirable and being disposed to  $\phi$ , a connection of much the same kind as that which exists between the belief that  $\phi$ -ing is good and the disposition to  $\phi$ . The 'normal' merely allows that it is a remote possibility that beliefs about desirability and desire come apart; a possibility that must be admitted because we must admit that subjects have only fallible access to their desires. But in order for this to provide a problem the objector would have to insist that there is no remote possibility that a subject who believes that  $\phi$ -ing is good, be not disposed to  $\phi$ . Yet what could the evidence for this claim be? Indeed, as far as evidence one way or the other is concerned, the evidence seems to point in just the opposite direction: see, for instance, Michael Stocker, 'Desiring the Bad' in *The Journal of Philosophy*, 1979, in which he reminds us that, in certain fits of depression, or self-deception, or in certain conditions of physical tiredness, one may believe that a certain course of action is good and yet be totally indifferent to it. I would myself resist the externalist conclusions that some would have us draw by consideration of such examples; that is, I resist the idea that the connection between moral judgement and motivation is wholly contingent, requiring some wholly added on desire to be moral. By all means, I say, let's give the internalist in ethics his due. But his due does not require any more than that we admit that there is a necessary connection of sorts between making a moral judgement and being disposed to act; his due requires only that we admit that, under certain conditions, moral judgement entails a disposition to act. But that we have admitted.

hybrid kind, over and above beliefs and desires, in order to explain why beliefs about the desirability of certain courses of action are, in a certain respect, practical. Nor, if this is right, do I see why we need to introduce a state of some further, mysterious, hybrid kind in order to explain why moral beliefs are, in a certain respect, practical. Rather, we should think of moral properties on the model of the property of being *prima facie* desirable. For we may then think of moral quasi-beliefs as being genuine *beliefs* about the properties that persons, actions, states of affairs, and the like have. But, since they have these properties in virtue of standing in certain relations to the desires that the subject has, they are beliefs that the subject can have only if he has certain *desires*.<sup>45</sup>

### 8. Nagel, McDowell and Platts again

If some version of the argument given in the last section is correct then we should be able to diagnose the flaws in the as yet unconsidered objections of Nagel, McDowell and Platts to the Humean theory; and, indeed, we should be able to diagnose the flaws in the theories of motivation with which they would replace the Humean theory. I want to close with some necessarily brief remarks aimed at showing that this is indeed the case.

(i) Early on in *The Possibility of Altruism* Tom Nagel puts forward an objection to the Humean theory that seems aimed not so much at *refuting* the theory, as at *deflating* it, thereby opening up room for his own non-Humean view:

The assumption that a motivating desire underlies every intentional action depends, I believe, on a confusion between two sorts of desires, motivated and unmotivated . . . The claim that a desire underlies every act is true only if desires are taken to include motivated and unmotivated desires, and it is true only in the sense that *whatever* may be the motivation for someone's intentional pursuit of a goal, it becomes in virtue of his pursuit *ipso facto* appropriate to ascribe to him a desire for that goal. But if the desire is a motivated one, the explanation of it will be the same as the explanation of his pursuit, and it is by no means obvious that a desire must enter into this further explanation. Although it will no doubt be generally admitted that some desires are motivated, the issue is whether another desire always lies behind the motivated one, or whether sometimes the motivation of the initial desire involves no reference to another, unmotivated desire.<sup>46</sup>

But there is confusion here only if the idea of a desire motivated by a state that is not a desire is itself plausible. Is it plausible? The Humean thinks not. He argues as follows.

<sup>45</sup> I have defended this view in my 'Should We Believe in Emotivism?' in *Fact, Science and Morality: Essays on A. J. Ayer's Language, Truth and Logic*, edited by Crispin Wright and Graham Macdonald, Blackwells, Oxford, 1986. See also my comments on McDowell in section 8 (iii).

<sup>46</sup> *The Possibility of Altruism*, p. 29.



A motivated desire is a desire had for a reason; that is, a desire the having of which furthers some goal that the agent has. The agent's having this goal *is*, in turn, *inter alia*, the state that constitutes the motivating reason that he has for having the desire (from (1)). But if the state that motivates the desire is itself a reason, and the having of this reason is itself constituted by his having a goal, then, given that the having of a goal is a state with which the world must fit rather than *vice versa* (from (2)), so it follows (from (3)) that the state that motivates the desire must itself be a desire. Thus, the Humean will say, the idea that there may be a state that motivates a desire, but which is not itself a desire, is simply implausible.<sup>47</sup>

The same point can be made in another way, by considering how, according to Nagel's own preferred theory of motivation, a state that is not a desire may yet motivate a desire. Suppose a subject now accepts the judgement 'At *t* I have (tenselessly: i.e. have or will have) reason to speak Italian'. This judgement may properly be thought of as the expression of a *belief* that the subject presently has. In Nagel's view, this belief may motivate a subject to promote his speaking Italian at *t* quite independently of the relation he believes *t* stands in to *now*.<sup>48</sup> Here, then, we have an example of a belief that, according to Nagel's theory, may suffice to explain a subject's action. And since whenever there is something sufficient to explain motivation there is enough to warrant the ascription of a desire, so, *a fortiori*, according to Nagel, that belief may also explain his having a desire; the desire to promote his speaking Italian at *t*. But, as is perhaps already evident, the Humean has a principled objection to Nagel's theory of motivation. For consider Nagel's theory in the light of (1)—the claim that having a motivating reason *is*, *inter alia*, having a goal. Does the subject who now believes that he has at *t* (tenselessly) reason to speak Italian necessarily now have a goal that he speaks Italian at *t*? Evidently not. The subject may indeed now believe that at *t* he has (tenselessly) a goal that he speaks Italian. But this belief cannot constitute his having a goal *now* that at *t* he speaks Italian, for it is a state with the wrong direction of fit (from (2)). Nor does the truth of this belief require that he has as a goal *now* that he speaks Italian at *t*. Rather, its truth requires that *at t* he has a goal that he speaks Italian.

The Humean will therefore say this about Nagel's theory of motivation, and hence about the theory according to which it is possible for a desire to be motivated by a state that is not itself a desire. Either this theory requires a conception of motivating reasons that is inconsistent with (1)—a claim we should have thought to be a conceptual truth about the connection between having a motivating reason and having a goal—or it requires a conception of belief that permits tenseless beliefs to be such that the world must fit with

<sup>47</sup> The objection here is that the idea that there is such a state is implausible. If we had some argument to the effect that the very concept of a state having 'both directions of fit' (see section 7) is incoherent, then we would be able to make a correspondingly stronger objection to Nagel's argument.

<sup>48</sup> *The Possibility of Altruism*, pp. 68–9.

them, that is, a conception inconsistent with (3)—a claim we should have thought to be a conceptual truth about the nature of desire.

This is a pleasant result. For it is on the basis of the distinction between motivated and unmotivated desires, and the alleged possibility that a desire may be motivated though not by another desire that Nagel, and McDowell following him, think that they can concede P<sub>2</sub> to the Humean—the claim that motivation requires the *presence* of a desire and means-end belief—without thereby accepting P<sub>1</sub>—the claim that motivating reasons are *constituted* by desires and means-end beliefs. But, if the argument given here has been correct, it emerges that P<sub>2</sub> is not a principle that we have any reason to believe in its own right. Rather, properly understood, our only reason for believing P<sub>2</sub> is that it is entailed by P<sub>1</sub>.

(ii) As I said earlier, Mark Platts in fact considers and rejects, a defence of the Humean theory based on a conception of desires as states with which the world must fit (see section 5). But why does he reject it? Platts rightly points out that we should demand of the Humean an ‘argument for the claim that any full specification of a reason for action must make reference to a mental state of the second, non-cognitive kind *vis-à-vis* direction of fit with the world’.<sup>49</sup> But he fails even to put a response to this demand in the Humean’s mouth. Rather, he asks rhetorically, ‘Why should it not just be a brute fact about moral facts that, without any such further element entering, their clear perception does provide sufficient grounding for action?’<sup>50</sup> But the answer to this can now be seen to be relatively simple.

Given that reason explanations are teleological, a subject’s clear perception of some moral fact could provide sufficient grounding for action only if (from (1)) it constituted his having some goal. But that would require in turn (from (2)) that the perception be a state with which the world must fit. However, a perception is not a state with which the world must fit. Thus we have reason to reject Platts’s suggestion. This is not, of course, to deny that the ‘clear perception’ of some moral fact may be something that a subject can have only if he has certain desires. This is parallel to the way in which the clear perception of the desirability of some state of affairs is something that a subject can have, in normal circumstances, only if he has certain desires. But then, though a subject who has such a clear perception may be disposed to act in certain ways, the state that grounds this disposition will be a desire, not a clear moral perception. I am therefore unimpressed by Platts’s own reasons for rejecting a defence of the Humean theory based on direction of fit considerations.

(iii) In short, this is my objection to John McDowell’s view that the virtuous agent may be motivated by his conception of a situation in which he perceives some moral requirement—something that McDowell claims is

<sup>49</sup> *Ways of Meaning*, p. 258.

<sup>50</sup> *Ibid.*

properly thought of as a cognitive state.<sup>51</sup> McDowell simply overlooks the possibility that, though the virtuous agent's conceptions are cognitive states, and though his having those conceptions suffices for his being in a state that motivates him, none the less the state that motivates him is a desire. This possibility is opened up once we realise that the Humean may admit that there are cognitive states that a subject can be in only if he is in some non-cognitive state.

(Nor need this be interpreted as the simple-minded idea that for each virtue an agent has, he has a corresponding desire, and that the condition for his having the conceptions associated with having that virtue is his having that desire.<sup>52</sup> Rather, we may think that having a virtue requires being disposed to have different sorts of desires in different sorts of situations. Taking this view allows us to hold that, though having certain desires is the condition for having the conceptions that a particular virtue makes possible, there is no straightforward mapping of desires onto virtues.)

## 9. Conclusion

It is, I hope, beginning to look as though the Humean theory is more plausible than many people have thought. The argument I have given for the theory has been really quite simple. However, it seems to have been overlooked by non-Humean theorists of motivation. Some have done so because they fail to distinguish motivating reasons from normative reasons. Others have done so because they hold a weak phenomenological, and hence inadequate, conception of desire. Most have done so because they have overlooked the implications of the fact that having a motivating reason is, *inter alia*, having a goal. However, once we keep it firmly before our minds that the Humean's is a theory of motivating reasons and equip ourselves with an adequate conception of desire, we see that only the Humean's claim that motivating reasons are constituted, *inter alia*, by desires is able to make proper sense of reason explanations as teleological explanations. For only an agent's desires may constitute his having certain goals. So, at any rate, I have argued.<sup>53</sup>

Department of Philosophy  
Princeton University  
Princeton, N.J. 08544  
U.S.A.

MICHAEL SMITH

<sup>51</sup> 'Are Moral Requirements Hypothetical Imperatives?', p. 18; 'Virtue and Reason', pp. 335-6, 345-7.

<sup>52</sup> 'Non-cognitivism and Rule Following', pp. 144-5.

<sup>53</sup> I am grateful to the following people for comments on an earlier version of this paper: Simon Blackburn, Jonathan Dancy, Gilbert Harman, Frank Jackson, Mark Johnston, Lloyd Humberstone, David McNaughton, Michaelis Michael, Nathan Tawil, and Jay Wallace.