

The Irrelevance of Direction of Fit

Iain Law

University of Birmingham
Edgbaston
Birmingham
B15 2TT

i.law@bham.ac.uk

The Irrelevance of Direction of Fit

The so-called 'Humean' view of motivation is pretty standard in the Philosophy of Mind. Its most prominent contemporary defender, Michael Smith, calls it a 'dogma'. Humeans believe in a strict divide between beliefs and desires. Beliefs have no intrinsic motivating force: I may believe anything at all, but only with the contribution of a separate desire will I be motivated to act. This claim should be broadened out to include all cognitive states (belief, knowledge...). The Humean claim is that cognitive states are wholly lacking in conative power. If some beliefs seem to us to motivate action, that can only be due to a contingent association with a separate conative state.

Humeans don't deny that such contingent connections between belief and motivation occur. They're very commonplace: if you were to acquire the belief that the building you're in is on fire, you'd immediately be motivated to do various things: to call the fire service, to warn others, to leave the building. Humeans are happy to accept contingent connections between belief and motivation. These connections may be sufficiently powerful and immediate that it would be very difficult for an agent to break the connection. This poses no problem for the Humean, since they merely want to insist that the belief is not necessarily linked to the motivation; that for an agent to be motivated there must be a separate mental state present: a desire.

One way to argue on behalf of the Humeans is to appeal to the 'direction of fit' of beliefs and desires¹. Both beliefs and desires may 'fit' the world, or not. If I have a false belief, my belief fails to fit the world. If I have an unfulfilled desire, there is a similar lack of fit. Beliefs and desires differ, it is suggested, in the 'direction' in which this 'fitting' relation operates. If my belief is false, it is the belief that fails to fit the world. If my desire is unfulfilled, it is the world that fails to fit with my desire.

The argument is that:

1. Cognitive states have the first direction of fit.
2. Motivational states have the second direction of fit.
3. No single mental state could have both directions of fit.
4. So no cognitive state can be a motivational state.

Resistance to this argument tends to focus on the third of these premises. Why we may not conceive of mental states that are at once both cognitive and conative? One well-known attack on this premise is M.O. Little's. Her elegant argument relies on the thought that a single mental state may have two contents, each with a different direction of fit.² My argument in this paper does not. I will explain how it can be that a mental state with a single propositional content can have both directions of fit, something that Smith says is 'just plain incoherent'³. The way in which Humeans defend the premise differs depending on whether they wish to adopt a normative notion of 'direction of fit', or a non-normative one.

¹ This argument can be found in Smith, 1994, 111-125

² Little, 1997, 64

³ Smith, 1994, 118.

In explaining the idea of direction of fit, it's hard to avoid using normative terms⁴. It's natural to say that the cognitive states *ought* to fit the world, but that this is not true of conative states (notice that it's a lot less comfortable to say that the world ought to fit conative states, since it is all too easy to think of desires that ought not to be satisfied – this is, I suggest, a significant fact which I'll be returning to).

Michael Smith is careful to avoid a normative construal of the distinction. This is because of the eventual purpose to which he wishes to put the Humean result, but that needn't bother us here.⁵ He instead characterises the distinction as 'a difference in the counterfactual dependence of a belief that p and a desire that p on a perception with the content that not p: a belief that p tends to go out of existence in the presence of a perception with the content that not p; whereas a desire that p tends to endure, disposing the subject in that state to bring it about that p.'⁶

Smith's non-normative version of the distinction doesn't work, since it fails to capture the intended difference. There are examples of both cognitive and conative states that fail to behave in the way that he describes. Think of the belief that you're going to fall from a high place: it can persist even though you perceive the solidity of the things holding you up. Psychologists and psychiatrists have documented many cases of delusional patients whose beliefs are resistant to correction⁷. There are examples of desires, too, which fail to persist as Smith says they tend to. 'Adaptive preferences' are desires that alter in the face of the perception that they will not be fulfilled. For example, Amartya Sen has reported on subjective perception of health in Indian men and women.⁸ Indian women typically rated their health as better than the men, despite the fact that they were no healthier from a third-person perspective. Sen suggests that their level of contentment with relatively poor health is a product of their knowledge that health care was less readily available to them than it was to the men. Their preferences have adapted to fit the realities of the situation: knowing that they will not receive good health care, their desire for it goes out of existence.

It may be possible to defend Smith's use of a non-normative notion of direction-of-fit as a way of separating beliefs and desires. For he says not that a belief will go out of existence in the presence of the perception of contradictory evidence, but that 'a belief that p *tends* to go out of existence in the presence of a perception with the content that not p; whereas a desire that p *tends* to endure...'⁹ Does this mean that my criticism of his account mis-fires?

This perhaps depends on how we are to understand the words 'tends to'. Obviously Smith cannot mean by this merely that usually beliefs behave in this way, but that

⁴ See for instance Mark Platts' well-known introduction of the term: '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 realization, and their realization is the world fitting with them; the fact that the indicative content of a desire is not realized in the world is not yet a failing *in the desire*, and not yet reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa.' Platts, 1979, 256-7

⁵ His reason, briefly, is that he wishes to employ these concepts in an analysis of normative judgement, and thus cannot appeal to normative notions for fear of circularity.

⁶ Smith, 1994, 115

⁷ For excellent discussion see Bortolotti 2005, 192-198

⁸ Sen, quoted in Nussbaum, 2000, 139

⁹ Smith, 1994, 115. Emphasis added.

sometimes they do not, or that most beliefs have this tendency, but that some do not. At least, he cannot mean this if we are to be able to take this as telling us what it is for a given mental state to be a belief rather than something else. Smith presents having direction of fit, understood in this dispositional way, as being definitive of belief, as capturing the essence of what it is to be a belief. That being the case, he cannot allow that some beliefs simply fail to have the feature in question. That would mean that there was something else that defined belief: some other criterion for being a belief. If there was some other criterion, of course he could then claim that most beliefs go out of existence in the face of contradictory perceptions. But the claim would then have lost its interest for us.

A better guess as to what Smith means is that each and every belief has a tendency, i.e. a disposition, to go out of existence in the presence of a contradictory perception. But just as other dispositions may sometimes fail to manifest themselves when something interferes with the conditions that normally produce the characteristic behaviour, so the disposition to go out of existence may be subject to interfering conditions. So, for example, sugar is soluble in water. This means that it has a disposition: to dissolve when immersed in water. But if a particular sample of sugar should fail to dissolve in water, that does not mean that it isn't sugar – it means that something interfered with the manifestation of its disposition to dissolve. Perhaps the water was already saturated with sugar, for example. Likewise, although a particular belief fails to go out of existence when I have a perception with contradictory content, this does not necessarily mean that my belief lacked the relevant disposition. It may instead mean that in this case something prevents this disposition from being manifested, even though it is present. Thus Smith can still claim that each and every belief has the relevant disposition. That doesn't commit him to saying that there are no circumstances under which it fails to be manifested.

I don't think that this response works for Smith. There are examples of beliefs to which we have no reason to ascribe the relevant tendency. Suppose that I wish to illustrate the Muller-Lyer illusion to my students. I carefully draw on the board two parallel lines of equal length, using a metre-stick. I pause to check that they are the same length, and that they look the same length. Then I add the necessary lines at the ends of the parallel lines. Stepping back and looking again, the parallel lines now seem to me to be different in length. But I still believe that they are the same length. My belief that they are the same length persists in existence despite the fact that I have a perception that they are not the same length. It's the same for other visual illusions. If I look at a version of the 'rotating snakes' illusion, it seems to me that some parts of the image are in motion in relation to other parts. But I persist in my belief that no part of the image is in motion. It seems that my beliefs about those lines and about that image, do not have the tendency Smith claims that all beliefs have.

So much for Smith's attempt to cash out the notion of direction-of-fit in non-normative terms. But what of a normative account of direction-of-fit? Isn't such an account made more plausible by the fact that all the counter-examples to Smith that were just mentioned are examples of beliefs and desires that are in some way defective (or, in the illusion cases, perceptions that are defective)? This suggests that even if we cannot say for all beliefs that they are states that will go out of existence in the face of contrary perception, we can perhaps say that they are states that *ought to* go out of existence in the face of contrary perception. But even this is not true, as the

case of visual illusions proves. Take the illusion that straight lines on a background of curved lines are themselves curved. After measuring, I believe that they are straight, and I ought to continue to believe this even though they continue to look curved. This suggests that beliefs ought to go out of existence only in the face of *veridical* contrary perceptions, or in the face of contrary perceptions that we have no independent reason to doubt. This seems right to me, but it means nothing more than the familiar thought that we should proportion our beliefs to the evidence.

Desires, on the other hand, are obviously not related to perceptions or to evidence like this. If I want to have cake, a perception that I do not have any doesn't mean that I ought to stop wanting some. Does this mean that the Humean argument is successful? I do not think so. We have identified a difference between being-a-belief and being-a-desire: established that these are distinct properties, but we have not yet shown that it is impossible for a single mental state to share these properties. It may look as though that is what has been shown, but only if we are careless. We need to be clear about what sorts of norms govern the two kinds of state.

For the Humean argument to work, the same sort of norm must apply on both sides of the distinction. For it is perfectly possible for a single state to exhibit both directions of fit if that state's reason to change and reason to remain unchanged are different sorts of reason. Here's an example: I will give you a vast sum of money if you believe that pixies are marching along the nearest wall. This is an instance of the familiar story in which some great reward will come to you if you can succeed in believing that *p* (*p* being something that you do not at present believe and indeed have evidence against). Suppose that you succeed – what shall we say of your belief that *p*? It is a state that you have epistemic reason to change, but one which you have prudential reason to maintain. You both ought to change it and ought to keep it, but this is no contradiction since these are two entirely different sorts of 'ought'.

Recall a comment I made above: it doesn't seem correct to say that a desire ought to persist in the face of a perception that the desire is not currently satisfied. Surely discomfort with saying this is rooted in disquiet about the content of the desire. Before we're prepared to say whether it should persist, we want to know what it is a desire for. What desires a person ought to have is a matter of prudential and moral norms, not epistemic ones.¹⁰ This means that there is no conflict between having the two directions of fit: it is simply a product of beliefs and desires being governed by different norms. There's no reason to think that a unitary state cannot be subject to two different sets of norms that pull in different directions.

This means that there is no reason to accept premise three of the Humean's argument.¹¹ To say that beliefs and desires have different directions of fit is to say no more than that they are subject to different sorts of norms in different sorts of way. In

¹⁰ This ceases to be so if we consider the indirect application of norms. So, epistemic norms may make it valuable to want to know things, because such a desire serves epistemic goals. Interestingly, this reveals the way in which even a single kind of norm may pull in two ways: imagine a belief that it is useful for an enquirer to have (i.e. it helps her to acquire knowledge) even though there is no evidence for the belief, or even though it is in fact false. In such a case, she has direct epistemic reason to lose the belief, and indirect reason to keep it.

¹¹ We can reject premise three without needing to make Little's move of claiming that a single mental state can have two contents, and thus two directions of fit. I have shown that even a single content can have two directions of fit, since a single content can be subject to more than one set of norms.

typical cases, beliefs are subject to epistemic norms, and desires are subject to moral and prudential norms. That this is true is irrelevant to whether or not Humeanism about motivation is true.

References

- Bortolotti, L. 2005 'Delusions and the Background of Rationality', *Mind & Language*, vol. 20.
- Little, M.O. 'Virtue as Knowledge: Objections from the Philosophy of Mind', *Nous*, vol. 31.
- Nussbaum, M. C. 2000 'Adaptive Preferences and Women's Options' in Nussbaum, *Women and Human Development*, (Cambridge: CUP)
- Platts, M. 1979 *Ways of Meaning*, (London: Routledge & Kegan Paul)
- Smith, M. 1994 *The Moral Problem*, (Oxford: Blackwell)