

The definitive version of this paper can be found in:

Ethical Theory and Moral Practice 2012.

Changing Direction on Direction of Fit

Alex Gregory

Abstract: In this paper, I show that we should understand the *direction of fit* of beliefs and desires in normative terms. After rehearsing a standard objection to Smith's analysis of direction of fit, I raise a similar problem for Humberstone's analysis. I go on to offer my own account, according to which the difference between beliefs and desires is determined by the normative relations such states stand in. I argue that beliefs are states which we have reason to change in light of the world, whereas desires are states that give us reason to change the world. I end by briefly discussing the relevance of my view to the Humean theory of motivation.

Perhaps the two most central mental states that humans have are beliefs and desires. Direction of fit accounts aim to say something about what these mental states are by drawing a certain contrast between them. Beliefs aim to fit the world, whereas desires aim to have the world fit them. But this metaphorical account is imprecise: we need some analysis of what, exactly, a direction of fit amounts to.

In this paper I argue against Michael Smith's (1987, 1994) and Lloyd Humberstone's (1992) attempts at this task, and try to put something in their place. In section 1, I briefly explain the nature and point of the direction of fit account. I then explain why Smith's account fails, and then argue that Humberstone's account is vulnerable to a similar objection. In section 2, I present my own view, according to which we should understand direction of fit in normative terms. In section 3, I briefly discuss the relationship between my view and the Humean theory of motivation.

(1)

There seems to be a certain contrast that we can draw between certain mental states that we have (Anscombe 2000: 56, Platts 1979: 256-7). Beliefs, and other belief-like attitudes, aim to have their content fit the world. They “inhale” content from the world. Another way to say this is that when the content of a belief fails to match up to how reality is, it is up to the belief to change. Beliefs contrast with desires, and other desire-like attitudes. Desires, and other desire-like attitudes, aim to have the world fit their content. They “exhale” content into the world. Another way to say this is that when the content of a desire fails to match up to how reality is, it is up to the world to change.

Whilst this picture is appealing, it is vague and metaphorical. Beliefs cannot literally aim to do anything, and it is not at all clear whether the point is that beliefs *should* fit the world, that they *normally tend* to, or something else. One might say that the content of a belief is one that is *believed* to fit the world. But that is circular. Completely parallel remarks apply to desire.

Given these difficulties, one might be tempted to just give up on the distinction between two directions of fit altogether. But the alternative option, given the intuitive appeal of the account, is to try to elucidate a more literal interpretation of what it amounts to, one that is neither implausible nor vacuous. This is what both Michael Smith and Lloyd Humberstone have tried to do.

Smith argues that:

The difference between beliefs and desires in terms of direction of fit can be seen to amount to a difference in the functional roles of belief and desire [...] a belief that p tends to go out of existence in the presence of a perception with the content not p, whereas a desire that p tends to endure, disposing the subject in that state to bring it about that p. (1994: 115)¹

¹ Here, Smith talks about bringing it about that P, as though the truth value of a proposition can change over time. But a standard thought is that the truth value of a proposition cannot change over time. A more precise claim might be that desires dispose us to have P be true as a result of our own efforts. But the complications here are large, and so I shall set this issue aside, both here and for the other theories I discuss below. I thank Ralph Wedgwood for bringing this problem to my attention.

But the problem with this account becomes clear when we ask what is meant by a “perception”. If perceptions are belief-like states, then the account defines the difference between beliefs and desires in terms of how they relate to belief, and this seems unhelpful at best (Humberstone 1992: 63-5). On the other hand, if perceptions are not belief-like states – if we understand them as mere appearances – then the account is false. Many of our beliefs tend to endure despite conflicting appearances, because we know that the appearances can be misleading. For example, I am not at all disposed to believe that things are colourless when I wake in the darkness in the middle of the night, though they appear that way. The same point applies to desires. It follows from these considerations that the account cannot be both informative and true (see also Sobel and Copp 2001: 46-51).

In light of this worry, Humberstone has proposed his own rival explanation of what direction of fit amounts to. He characterises directions of fit in terms of background conditional intentions (1992: 73-5). On Humberstone’s view, whether a mental state is a belief or a desire is determined by our higher-order intentions regarding that state. On his view, a mental state that P is a desire that P only if the agent has a conditional intention that P, an intention conditional on their having that mental state. In contrast, a mental state that P is a belief that P only if the agent has a conditional intention that they cease to have that mental state, an intention conditional on P being false. More concisely, desires are states we intend to satisfy while we have them, beliefs are states we intend to abandon if false.

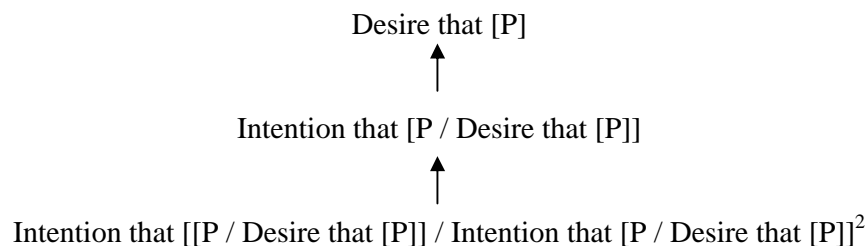
But just as the problems with Smith’s view became clear when we reflected on what he meant by a “perception”, so too do the problems with Humberstone’s account become clear when we reflect on what he means by an “intention”. Again, the objection comes in the form of a dilemma. (Actually, as he describes the view, I think that it falls to the second horn of the dilemma. But casting the objection in this broader manner will make clear that the worry is not merely about the precise details of the view.)

The dilemma relies on the claim that desires and intentions have the same direction of fit. This seems plausible. We describe intentions as fulfilled or unfulfilled, not true or false. And intentions are states that have a content (*what it is* that you intend to do), and that content is one that the world should change to fit, not the reverse: we should act so as to fulfil our intentions, not change our intentions so that they are already fulfilled.

If intentions have the desire-like direction of fit, then need some explanation of how this is so (cf. Humberstone 1992: 74-75). Here is the dilemma: either we try to use Humberstone's account to make sense of the direction of fit of intentions, including the higher-order intentions that govern beliefs and desires, or else we use some other account. I take the options in turn.

First, we might try to employ Humberstone's account to understand how intentions have the direction of fit that they do. In this case, we presumably analyse the direction of fit of intentions in the same way that we analysed the direction of fit of desires. We should say that an intention that P has the desire-like direction of fit because of a higher-order intention that [P given the presence of the intention that P]. But now we should ask what we should say about such higher-order intentions. They explain the direction of fit of beliefs, desires, and first-order intentions, so it is important that we have a good grasp on what they are supposed to be. Presumably they too have the desire-like direction of fit, so we should apply Humberstone's analysis again. A natural thought is that such higher-order intentions have their direction of fit in virtue of some still higher-order intention, in particular, the intention to satisfy the prior intention given its presence.

It is helpful to represent this view diagrammatically. I shall present the problem for the case of desire, but it should be clear the worry generalises. In the following diagram, “/” indicates conditionality, and the arrows represent explanations of the direction of fit of some state:



The desire that P has its direction of fit explained by the intention immediately below it, which in turn has its direction of fit explained by the intention below it. But, and this will be obvious by now, it is clear that this process will never stop. Each time we try to explain the

² There is a worry that the conditional operator is being embedded inappropriately here. If so, all the worse for Humberstone's view on this horn of the dilemma.

direction of fit of some state, we appeal to some state which itself has a direction of fit which needs explaining. This is problematic in part because it suggests that in order to have a single belief, desire, or intention, an agent needs to have an infinite numbers of intentions. More seriously, this is problematic because if this explanatory chain never comes to an end, we never really get an explanation of how the original state has the direction of fit that it does. So on this horn of the dilemma, Humberstone's view generates a vicious regress (cf. Zangwill 1998: 180).

In summary, on the first horn of the dilemma we claim that we should understand the direction of fit of intentions, including higher-order intentions, in the same way that Humberstone suggests that we understand the direction of fit of desire. But this generates a vicious regress, and so the account must be false. It follows that Humberstone should not opt for the first horn of the dilemma. What about the second horn? On this horn of the dilemma, we try to explain the direction of fit of intentions in some way that is independent of Humberstone's account. This view avoids the regress above.

But if we hold Humberstone's account in this way, then Humberstone's account is clearly incomplete. It needs supplementing with another direction of fit account to make sense of the direction of fit of intentions, and it is not clear why that other account could not itself be used to make sense of the direction of fit of beliefs and desires in the first place. Further, there is a worry that if Humberstone's account cannot explain what it is for an intention to have the desire-like direction of fit, it seems unsatisfying, at best, to appeal to the nature of such intentions to explain the direction of fit of beliefs and desires themselves.

So either we do use Humberstone's account to make sense of the direction of fit of intentions, or we do not. If we do, the account generates a vicious regress, and if we do not, the account is inadequate. Either way, the account is unappealing.

It seems that both Smith and Humberstone have provided direction of fit accounts that fail to fit the truth. In light of this, we had better revise not the truth, but instead our understanding of direction of fit. But how? Let us briefly consider the overall form of Smith's view. It is that we explain the direction of fit of beliefs and desires by reference to a different kind of mental state: perceptions. In this respect, it is similar to Humberstone's view. On Humberstone's

view, we explain the direction of fit of beliefs and desires by reference to a different kind of mental state: intentions. But in both cases the view fails, at least partly because the third mental state is itself best understood as having a certain direction of fit.

And that is precisely the problem. Many of our mental states seem to be either belief-like or desire-like in their direction of fit. That is why the direction of fit account is so exciting: it promises to explain at least part of the nature of many of our mental states. But if that is true, it seems unlikely that we shall be able to explain what we mean by a direction of fit in terms of some mental state. Our mental states are supposed to be the things getting explained, not the things doing the explaining (cf. Sobel and Copp 2001). We cannot, *pace* Smith and Humberstone, understand what it is to have a direction of fit by reference to the nature of some mental state. We will have to understand direction of fit in terms of something else.

(2)

The obvious way to go, especially given how the direction of fit metaphor was introduced – in terms of *mistakes* – is to understand direction of fit in normative terms (Anscombe 2000: 56, Platts 1979: 256-7, Zangwill 1998). In broad terms, we could say that we should change the world in the wake of our desires, and should change our beliefs in the wake of the world. The challenge is to understand these claims in a way that makes them plausible. I shall aim to precisely describe the normative requirements that govern beliefs and desires, and I shall show how such an account is not vulnerable to two objections raised against normative direction of fit accounts raised by David Sobel and David Copp (2001).

I take it that this broad move is appealing enough. Many have thought that it is definitive of beliefs that they should aim at the truth (e.g. Gibbard 2003, Shah 2003, Wedgwood 2002). Given how tempting the direction of fit picture is, it should hardly be surprising if some similar claim is true of desire: that they aim at satisfaction.³ In both cases, the basic claim is that beliefs and desires cannot be understood without reference to the norms that govern them. One might object that it is impossible to reach normative conclusions by merely studying the nature of beliefs and desires, on the ground that this breaches some kind of is/ought gap. But the objection misses the point: on the view being suggested, beliefs and desires are already at least partly on the ought side of this divide.

³ This should not be taken as ruling out the possibility that desires have other aims as well.

Enough of the generalities: how should we fill in the details of this account? An initial view might be this:

For a mental state that P to be a belief that P is for it to be the case that if not-P, you have a reason to abandon the state.

For a mental state that P to be a desire that P is for it to be the case that if not-P, you have a reason to make P the case.

This view is false. But before I say why, and modify the account in the face of those problems, let me note four merits of the account, even in this simplistic form. They will be retained by the more sophisticated version of the account I eventually endorse, so we may as well put them in the open now.

First, though I shall present the account just as one of beliefs and desires, I hope it is clear that the account could easily be expanded to make sense of other states with the belief-like or desire-like direction of fit, such as intentions. Intentions can equally be understood in terms of the reasons we have to make their contents true.⁴ But note that the account also does a good job of distinguishing beliefs and desires from other states which they are similar to, but with which they do not share a direction of fit, such as imaginings and compulsions. Imaginings, unlike beliefs, are not states which we have reason to constrain in the face of the world (cf. Velleman 1992: 111-114). Compulsions, unlike desires, are not states that we necessarily have reason to satisfy.

Second, the account says that we have reasons to abandon false beliefs. It does not say that we have reasons to adopt true beliefs. This seems to me to be correct. It is a failure *of belief* if it is false. But though an agent might suffer for failing to have a true belief, the failure on their part isn't a fault in any belief that they have. Such failures have little to do with the direction of fit of beliefs. Moreover, it would be a severe problem for any view which said that we did have reason to adopt all true beliefs. That is clearly false: I have no reason at all

⁴ This doesn't entail that there is no distinction between desires and intentions. It might be that the reasons intentions generate are of a different kind to – perhaps stronger than – those that desires generate, and it might be that intentions have other features that are independent of their direction of fit that distinguish them from desires.

to have a true belief about the number of blades of grass in your garden. That just isn't something I should worry about.

Third, the account only says that we always have a reason to abandon false beliefs, and that we always have a reason to satisfy our desires. It does not say that we always have a *decisive* reason to abandon false beliefs, or that we always have a *decisive* reason to satisfy our desires. This seems right. It is sometimes true that we have false beliefs that are too costly to abandon, all things considered. And it is very often true that we have desires that are too costly to satisfy, all things considered.

Fourth, the account states that we have reasons to satisfy our desires. It does not say that we have reasons to do what we believe will satisfy our desires (cf. Brandom 2000: 89-90). This might seem problematic. But note that the account is compatible with the idea that when I believe that something will lead to the satisfaction of a desire of mine, I should form an instrumental desire for that thing. The account then implies that this instrumental desire does change what I have subjective reason to do. More generally, it seems that we have reason to do what we believe will satisfy our desires *because* we have reasons to satisfy our desires. So the account rightly states the most fundamental normative constraint on desire.

Still, despite these merits, this account is indeed false. I shall offer two objections to the account, and modify it in the face of each.

(2.1) First objection: Reasons are too easy to come by

This objection applies to both the account of belief and the account of desire. Take beliefs first. I might have a mental state with the content P. And it might be the case that, if not-P, I have reason to abandon the state. But the state still might not be a belief. For instance, imagine that I intend that you will die a horrible death. Because one should never intend that others die horrible deaths, I have reason to abandon this state. It trivially follows that I have reason to abandon this state whenever its content is false. So this state meets the above conditions for belief, but isn't a belief (cf. Zangwill 1998: 177). A parallel objection applies in the case of desire. Assume that everyone always has reason to make it true that they are happy. It would then follow from the account above that *any* mental state with the content "I am happy" is a desire. But this is obviously false.

The problem that these examples raise is that the account is too indiscriminating as to the explanation of why we have the reason that we do. But we can modify the account to avoid such objections:

For a mental state that P to be a belief that P is for it to be the case that if not-P, *the fact that not-P gives you* a reason to abandon the state.

For a mental state that P to be a desire that P is for it to be the case that if not-P, *the mental state gives you* a reason to make P the case.

This modified account avoids the above two objections. With respect to nasty intentions, what gives you reason to abandon such states is not simply that their content is false.⁵ And with respect to mental states whose contents are things that we have reasons to make true, only those states which *provide* reasons are desires.

Note that this revised account is ambiguous. The phrase “gives you a reason” could mean “is a reason” or could mean “explains why you have a reason”. Deciding between these alternatives would require a much broader investigation into what reasons are, an investigation further complicated by the modifications I make to the view in the next section. So I shall here stay neutral between these alternatives, and stick with the ambiguous formulation.

In summary, the account should appeal not only to the normative relations in which beliefs and desires stand, but also the grounds of those normative relations. Beliefs are states which the world can give us reasons to change, and desires are states which give us reasons to change the world.

(2.2) Second objection: Bad desires

This begins as an objection to the account of desire, but we shall see that it develops into a problem for belief as well. Imagine that I desire to be Prime Minister. According to the account above, this desire gives me a reason to be Prime Minister. But this looks like it could

⁵ Note that the same point applies to desires or intentions that we should abandon because they are too difficult to satisfy. What gives us reason to abandon such states is not simply that their content is false, but also that it will remain so.

well be false. Perhaps I am lazy, ignorant and malicious, and so have no real reason to be Prime Minister at all.

The response to this seems simple. We modify the account of desire:

For a mental state that P to be a desire that P is for it to be the case that if not-P, the mental state gives you a *subjective* reason to make P the case.

According to this account, my desire to be Prime Minister does not give me *objective* reason to be Prime Minister. Perhaps that is a terrible possibility, with no real justification at all. But nonetheless, *from my point of view*, something speaks in favour of my being Prime Minister: my desire can rationalise that course of action for me. So I do have *subjective* reason to be Prime Minister.

I think that this response to the objection is sound. But it raises two further problems.

First, it raises the problem of understanding the distinction between subjective and objective reasons. It would take us too far from the present task to engage in that task in detail. At any rate, I can afford to be fairly ecumenical. I take it that everyone agrees that there is *some* sense in which an agent is subject both to requirements based on how things really are, and also requirements based only on features of the agent.⁶ So long as there is a distinction along these lines, I can appeal to it in order to formulate my account. It matters little exactly how the distinction is drawn.

The best way to see this point is to think about intentions. We are subjectively, but not objectively, required to fulfil our intentions. If I intend that P, then I am subjectively required to bring it about that P, and subjectively required to do so even if I am objectively required *not* to bring it about that P. Desires make a difference of a similar kind, though less strictly. If I desire that P it need not be true that I am subjectively *required* to bring it about that P, but certainly something subjectively counts in favour of my bringing it about that P, and this is

⁶ Of course, that P will make me feel happy might generate an objective reason for me to bring it about that P. So it is true that sometimes features of the agent generate objective, rather than subjective, reasons. The precise contrast is between requirements based on features of the agent that are themselves normatively assessable, and requirements based on anything else.

true regardless of whether anything objectively counts in favour of my bringing it about that P.

Though I think this is broadly correct, it is unfortunate that two of the major accounts of the subjective/objective distinction sit unhappily with my view. So let me briefly take the two accounts in turn, explain why they sit unhappily with my view of desire, and then explain why this is a problem for them and not for my view.

The first popular account of this distinction is the wide scope account. This view says that we should understand subjective requirements as objective requirements to have/perform certain *combinations* of attitudes/actions (e.g. Broome 2004). On this view, for some attitude of mine to subjectively require me to bring it about that P is for me to be objectively required to either [abandon this attitude or bring it about that P]. As usually stated, the wide scope account is understood in terms of oughts rather than reasons. But it is easy to see that the account could be supplemented in this respect. We could say that for some attitude of mine to give me a subjective reason to bring it about that P is for me to have an objective reason to either [abandon the attitude or bring it about that P].

In this form, the wide scope account is problematic for my account of desire. It understands subjective requirements in a *symmetric* manner. It says that if some attitude of mine subjectively requires me to bring it about that P, then I can meet this requirement either by abandoning the attitude or by bringing it about that P. But this former possibility is exactly the possibility that the direction of fit of desire is supposed to rule out. The direction of fit account should militate against the idea of abandoning a desire merely because its content is false. But my direction of fit account, when understood in wide scope terms, fails to have this implication.

Thankfully for my purposes, it is widely accepted that the wide scope account needs amending in this respect (e.g. Kolodny 2005: 528-530, and Broome *manuscript*). Consider the following subjective reason. If I believe that I ought to bring it about that P, I have a subjective reason to intend to bring it about that P. If the wide scope view were the whole story about subjective requirements, if you find yourself believing that you ought to bring it about that P but failing to intend that P, you can satisfy the above requirement by abandoning

your belief that you ought to bring it about that P. But changing one's normative beliefs to fit one's prior intentions is not a rational thing to do. So the wide scope account needs supplementing with some principles about *how* we should satisfy wide scope requirements. And such supplementary principles would allow us to understand my claim about desire in a way that makes clear that unsatisfied desires should generally be satisfied, not abandoned. In summary, the wide scope account fits badly with my account of desire, but only because of a feature it has which ought to be changed for independent reasons.

The second popular account of the distinction between subjective and objective reasons is a view often confusingly named the "subjective reasons account" (e.g. Schroeder 2009). This account says that to have a subjective reason to bring it about that P is to believe that Q and for Q, if it were true, to be a reason to bring it about that P. So for example, I have a subjective reason to fill my car up with fuel if I believe that my car needs fuel, and for the proposition that my car needs fuel, if it were true, to be a reason to fill my car up with fuel.

This account also fits badly with my account of desire. It entails that only our beliefs can generate subjective reasons. But this seems unduly restrictive: it is clear that intentions can generate subjective reasons, and appealing to think that desires do as well (cf. Setiya 2007). One possibility here is to say that both desires and intentions must always be accompanied by certain beliefs, so that desires and intentions generate subjective reasons because of their accompanying beliefs (cf. Schroeder 2009). Either this is plausible, in which case this account is compatible with my account of desire, or else this is implausible, and we shall have to supplement this account in order to explain how mental states other than beliefs generate subjective requirements, and this new account could be used to make sense of my account of desire. I need not decide between these options here.

In short, I think we should be happy with the idea that desires could be defined in terms of the subjective reasons they provide, no matter what account of subjective reasons we should endorse.

But there is a second objection to modifying the account of desire so that it refers to subjective, rather than objective, reasons. It is that if we modify the account of desire in this way, we should do the same for belief. We should say that:

For a mental state that P to be a belief that P is for it to be the case that if not-P, the fact that not-P gives you a *subjective* reason to abandon the state.

But this account of belief seems false, for two reasons. First, it seems like an *objective* failing in a belief that it is false. Nothing you could learn about the agent makes a false belief any more successful as a belief. Second, subjective reasons are supposed to be reasons generated by features of the agent. But in this case the reason is not provided by a feature of the agent at all, but instead by a feature of the world.

The solution to this problem is just to deny that we must modify the account of belief to match the modified account of desire. We should hold a hybrid view:

For a mental state that P to be a belief that P is for it to be the case that if not-P, the fact that not-P gives you an objective reason to abandon the state.

For a mental state that P to be a desire that P is for it to be the case that if not-P, the mental state gives you a subjective reason to make P the case.

At this stage, some will object that I no longer have something that resembles the original direction of fit account. That account described a contrast between beliefs and desires such that they were directly opposed to one another. But this new account no longer has that feature. It says that the contrast not just in their *direction* of fit, but also in the *kind* of fit being sought. Beliefs should objectively fit the world, whereas desires should subjectively have the world fit them. It's not clear that this is the contrast the direction of fit account was supposed to capture.

In fact, I think this worry is mistaken. Not merely because the true direction of fit account *could* have this implication, but because the true account *must* do so. One way of describing the direction of fit contrast is that beliefs have the world-to-mind direction of fit, whereas desires have the mind-to-world direction of fit. I have avoided using these terms for the simple reason that it is hard to keep in mind that “world-to-mind” means “world to influence mind” and not “world to fit mind”, and vice versa. But this way of talking does have one advantage: it shows that beliefs and desires contrast *both* in what their direction of fit runs *from* and in what their direction of fit runs *to*. The direction of fit of belief is responsive to the world, and encourages changes in the belief. The direction of fit of desire is responsive to the

desire, and encourages changes in the world. These states *do* contrast in two ways (cf. Anscombe, who talks only about the *location* of the mistake (2000: 56), as though the *explanation* of the mistake weren't also at issue (cf. Zangwill 1998: 177)).

Because the direction of fit of desire runs from the desire, it is hardly surprising that the reasons generated are subjective. They are reasons generated by the agent. In contrast, the direction of fit of belief runs from the world, and so it is hardly surprising that the reasons generated are objective. In addition to this, the reasons that desires generate suggest that we change the world. In contrast, the reasons the world generates suggest that we change our beliefs. Beliefs and desires contrast in both of these ways. That my account postulates a pair of contrasts is a good feature of the account, not an objection to it.

(2.3) Remaining Objections

Since this is the last modification I make to the view, the view I endorse is this:⁷

For a mental state that P to be a belief that P is for it to be the case that if not-P, the fact that not-P gives you an objective reason to abandon the state.

For a mental state that P to be a desire that P is for it to be the case that if not-P, the mental state gives you a subjective reason to make P the case.

This account is unashamedly normative. It says that the difference between the direction of fit of beliefs and desires is a difference between the reasons those states relate to. To determine whether some state has the direction of fit of belief or of desire, we have to determine what we have reason to do, and why, when the content of that state fails to match up to the facts. If the state gives us a reason of a certain kind to change the facts, then the state has the direction of fit of desire. If the facts give us a reason of a different kind to change the state, then the state has the direction of fit of belief. On this account, the contrast between directions of fit is a contrast between two normative relations in which mental states can stand.

⁷ One worry I have not addressed is that the account wrongly says that if I desire P, but desire not to desire P, I still have a subjective reason to bring it about that P. This seems acceptable to me: someone with those desires could more intelligibly bring it about that P than someone with neither desire. But even if this is wrong, I could simply modify the account of desire so that it says that we have a reason to satisfy a desire except when there is a second-order desire not to have it. I thank David McNaughton for this objection.

David Sobel and David Copp argue that no such normative account could be true (2001: 51-2). They have two objections:

The first objection comes in the form of a dilemma. Either the difference between the direction of fit of belief and the direction of fit of desire is understood in terms of their normative relations to the world, or else it is understood in terms of their normative relations to other mental states (for the latter, see Zangwill 1998, especially 194). In either case, they think the account is problematic. It might seem that my account falls onto the first horn of this dilemma, so I shall address their objection to this option.

On this horn of the dilemma, we define the difference between the direction of fit of belief and the direction of fit of desire in terms of their normative relations to the world. In this case, they write the following. I quote at length:

“A view of the former kind presumably would be some variation of the idea that the belief that p ought to be held only if p , whereas the desire that p can properly be held even if it is not the case that p . But although, of course, desires are not capable of truth, desires can be assessed as ‘satisfied’ or not. A view of this kind therefore presupposes that belief is unlike desire in that a belief is capable of truth whereas a desire is merely capable of being satisfied. As Zangwill remarks, this ‘just raises the problem of how to distinguish between truth and satisfaction’ (173). [...] This was the distinction that the metaphor of direction of fit was supposed to explicate rather than presuppose. We therefore set aside normative approaches of the first kind.”

(2001: 51)

I am not quite clear what this objection says. My best guess is that they are making two separate objections. The first is a worry that normative direction of fit accounts will mention a distinction between truth and satisfaction in their definitions of the normative relations that govern beliefs and desires. My account does no such thing, and so this is not a problem for my account. But I do not think that this is all of their worry, mainly because I do not see why this should be a problem that affects a normative direction of fit account more than any other direction of fit account.

So I wonder if they are also worried about whether the normative standards that govern belief and desire are actually different. In both cases, they might say, it is good if the content of the state is true. And in both cases, it is bad if the content of the state is false (cf. Zangwill 1998: 177-8, 194). So, they might think, there is no way to distinguish beliefs from desires in normative terms. But though it might be true that the kind of failing each state can have is similar, the difference lies in what we should do about it. I have said that in the face of the kind of badness where the content of a belief fails to fit the world, we have reason to change the belief. And in the face of the kind of badness where the content of a desire fails to fit the world, we have reason to change the world. So this objection fails as well.

In short, this objection poses no problem for my account. Their second objection is that the fundamental difference between the direction of fit of belief and the direction of fit of desire cannot be normative because all normative facts are explained by non-normative facts. So whatever the putative normative difference is between the direction of fit of belief and the direction of fit of desire, there must be some non-normative difference between the two states that *explains* this normative difference (2001: 52). But then it looks as though this non-normative difference is what really distinguishes the two states, and the normative difference between them is just a consequence of this. It follows that *no* normative direction of fit account could be true. There are two things to say about this objection.

First, it is not clear whether it is true that all normative facts are explained by non-normative facts. It is unclear whether there are non-normative facts that explain the normative facts that murders are wrong, that sharp knives are good knives, that you have reason to do what you ought to do, and so on. Perhaps facts about the direction of fit of belief and desire are like these other normative facts in that they have no non-normative explanation.

Second, it is possible to be more concessive. Perhaps they are right that any normative fact must be explained by some non-normative fact. That, in conjunction with the account offered here, would simply entail that beliefs and desires are not, at the fundamental level, distinguished by their respective directions of fit. This does not seem like an objectionable view, and holding this does not undermine the thought that there is theoretical interest in the direction of fit account.

Furthermore, it might be that the non-normative differences between beliefs and desires are too heterogeneous to be interesting. It might be true that we can see the interest in the distinction between beliefs and desires only once we look at these states at a greater level of abstraction, in terms of the commonality they have at the normative level. For comparison, think of the distinction between good and bad football players. Perhaps we could draw this distinction in wholly non-normative terms, but I doubt that the resulting distinction would look very interesting. Perhaps the relevant facts that we care about, with respect to football players, as well as with respect to beliefs and desires, are only the normative facts, even if they are not the most fundamental facts around.

(3)

The final thing to address is the issue to which the distinction between directions of fit has prime importance. That is Humeanism. My account of desire claims that desires are the states which provide us with subjective reasons to make their contents true. Can this justify the Humean theory of motivation? Exploring this in detail would take us far afield. But let me make two short suggestive remarks.

First, note that for some state to motivate an action, it needs to do more than to *bring about* the action. It has to also *rationalise* it (Davidson 2001: 3). If a state is a desire just so long as it provides us with subjective reasons to act, one can see why we might think that only desires can motivate actions. Only desires have the right kind of normative force to rationalise the events they generate in such a way as to make them actions. So the account I defend here might well imply that only desires can motivate us to act, might well justify the Humean theory of motivation.

But second, note that despite this, my account does not entail that desires are just dispositions to be motivated. On my account, one will be disposed to be motivated by one's desires only if one is disposed to be motivated to do what one has subjective reason to do. But this is sometimes false. It is most clearly false when an agent is irrational. For example, I might desire to finish this paper today, believe that I can do this only if I stop playing solitaire now, and yet fail to be motivated to stop playing solitaire now. So though my account will entail something similar to the account that Smith offers (1987, 1994), it nonetheless diverges in cases where agents are irrational. And this seems like a plausible result.

One final point. Though my account might justify something like the Humean theory of motivation, there is one complication. Some people will think that our beliefs about what we have reason to do provide us with subjective normative reasons to act. They will think that believing you have (objective) normative reason to bring it about that P is enough to provide you with a subjective normative reason to bring it about that P. If we accept my direction of fit account, there are two ways to deal with this suggestion. The first, appealing to Humeans, will be to claim that such people are mistaken. We might claim that such beliefs *cannot* in fact, provide us with subjective reasons to act. Such beliefs are like any other belief: a slave of the passions. The second possibility is perhaps more appealing. It is to claim that such beliefs *just are* desires. I hope to argue for this second view elsewhere.

Conclusion

In this paper I have done three things. First, for completeness, I rehearsed an objection to Smith's direction of fit account. Second, I argued that a similar problem besets Humberstone's direction of fit account, and I explained the common root of these problems. Third, I argued in favour of an alternative direction of fit account, one not subject to this problem, nor, so far as I can see, any other.

Beliefs and desires are to be understood in this way:

For a mental state that P to be a belief that P is for it to be the case that if not-P, the fact that not-P gives you an objective reason to abandon the state.

For a mental state that P to be a desire that P is for it to be the case that if not-P, the mental state gives you a subjective reason to make P the case.

I can only hope that this account fits the truth.⁸

Bibliography

Anscombe, G.E.M. *Intention* (2nd Ed.), Harvard (2000)

Brandom, R. *Articulating Reasons*. Harvard (2000)

⁸ I thank Jonathan Dancy, Federico Lauria, Bart Streumer, and the audience at the 2011 annual BSET conference, for many extremely helpful comments on this paper.

- Broome, J. 'Reasons' in Wallace, J., Pettit, P., Scheffler, S. and Smith, M. (Editors). *Reason and Value: Essays on the Moral Philosophy of Joseph Raz*. Oxford University Press (2004) pp.28-55
- Broome, J. *Rationality through Reasoning*, unfinished manuscript
- Davidson, D. 'Actions, Reasons, Causes' (2001a) in his (2001) pp.3-20
- Gibbard, A. 'Truth and correct belief' in *Philosophical Issues* 15 (2003) pp.338-50
- Humberstone, L. 'Directions of Fit' in *Mind*, Vol. 101, No. 401, (1992), pp.59-83
- Kolodny, N. 'Why be Rational?' in *Mind* 114:455 (2005) pp.509-563
- Lewis, D. 'Probabilities of Conditionals and Conditional Probabilities' in *The Philosophical Review* 85:3 (1976), pp.297-315
- Platts, M. *Ways of Meaning*, Routledge (1979)
- Schroeder, M. 'Means-end coherence, stringency, and subjective reasons' in *Philosophical Studies* 143:2 (2009) pp.223–248
- Shah, N. "How Truth Governs Belief" in *The Philosophical Review* 112:4 (2003) pp. 447-482
- Setiya, K. 'Cognitivism about Instrumental Reason' in *Ethics* 117:4 (2007) pp.649-673
- Smith, M. 'The Humean Theory of Motivation' in *Mind*, Vol. 96, No. 381, (1987), pp. 36-61
- Smith, M. *The Moral Problem*, Blackwell (1994)
- Sobel, D. and Copp, D. 'Against Directions of Fit' in *Analysis* Vol. 61, No. 1, (2001), pp. 44-53
- Velleman, D. 'The Guise of the Good' in *Nous* 26 (1992) pp.3-26
- Wedgwood, R. 'The aim of belief' in *Philosophical Perspectives* 16 (2002) pp.267-97
- Zangwill, N. 'Direction of fit and Normative Functionalism' in *Philosophical Studies* Vol. 91, (1998) pp.173–203