

Against direction of fit accounts of belief and desire

DAVID SOBEL & DAVID COPP

Champions of Humean belief–desire psychology have long tried to flesh out Hume’s claim that a desire does not contain ‘any representative quality, which renders it a copy of any other existence or modification’ (Hume

[1739–40] 1978: 416). Hume surely meant to be contrasting desires with beliefs, which do have a representative quality and, if true, do copy or mimic the world as it is. The most influential modern version of Hume's claim is that desires have a different 'direction of fit' than beliefs. Beliefs 'aim to track the world', while desires 'aim to impose themselves onto the world'.

Elizabeth Anscombe offered a useful anecdote to help us understand direction of fit. She wrote of a man shopping with a list, and a detective trying to write down what the man buys. She said that

if the list and the things that the man buys do not agree, and if this and this alone constitutes a *mistake*, then the mistake is not in the list but in the man's performance ... whereas if the detective's record and what the man actually buys do not agree, then the mistake is in the record. (1957: 56)

Michael Smith (1994) laudably attempts to get beyond metaphors to a more precise presentation of the root thought. He writes

For the difference between beliefs and desires in terms of directions of fit can be seen to amount to a difference in the functional roles of belief and desire. Very roughly, and simplifying somewhat, it amounts, *inter alia*, to 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*. (1994: 115)

This, we think, is an admirable expression of what modern champions of the Humean thought have been getting at with their metaphorical talk of directions of fit. However, we claim that Smith's attempt to make this notion more precise is open to general objections that cast real doubt on the helpfulness of talk of directions of fit in understanding beliefs and desires.

The direction of fit understanding of beliefs and desires can be given a descriptive or a normative interpretation. One might claim, as Smith does, that beliefs and desires actually do play different functional roles and can be distinguished descriptively. Alternatively, one might claim that beliefs and desires ought to play different functional roles, that they are subject to different normative requirements. For example, one might claim that, in Mark Platts's words, 'beliefs should be changed to fit with the world, not vice versa' while 'the world, crudely, should be changed to fit with our desires, not vice versa' (1979: 257, as quoted in Smith 1994: 112 – Platts does not actually endorse this view). Anscombe's understanding of the difference in direction of fit appears to have been normative. We initially focus

on Smith's descriptive formulation. At the end of the paper, we return briefly to normative variants.

1. *Problems for Smith's introduced state*

Smith's account distinguishes between a background state of the agent (which might be either a belief that p or a desire that p) and an introduced state (the perception with the content that not p). His thesis is that what makes a background state a belief or a desire is its tendency to remain or be driven out upon the introduction of the introduced state. One might worry that Smith's approach involves a departure from the basic idea of direction of fit theory, which is that the distinction between belief and desire turns on the different ways that their existence depends on the *facts*. But there is no reliable tendency for our beliefs to go out of existence in the face of contrary facts unless the facts are 'taken in' cognitively. For this reason Smith, and advocates of similar accounts, need to introduce a state that represents the facts in order to get the world into cognitive space.

An account of what makes a background state be a belief or a desire does not take the notions of belief and desire as primitives. Rather, the aspiration is to explain what beliefs and desires consist in.¹ Smith is offering us a 'conception' of desire that he intends as an answer to the question, 'What is a desire?' (104). He tells us that on his conception, 'desires are states that have a certain functional role' (113), a role that contrasts with the functional role of beliefs. The passage on directions of fit is meant to explain in part what the two functional roles are. For this reason it is crucial that the introduced state is not presupposed to be a belief or a desire, and that Smith's account not presuppose the concept of belief or desire in any way, for this would make the explanation unhelpfully circular. This is surely why Smith specifies that the introduced state is 'a perception with the content that not p ' rather than the circular 'the belief that not p '. But why does he specify 'a perception with the content that not p ' rather than, more simply, 'a perception that not p '?

To describe a person as perceiving that not p , or as having a perception that not p , implies that the person believes that not p . If we learned that a person did not believe that the road is wet we would not describe her as 'perceiving' that it is wet. It would therefore be circular to take the intro-

¹ In attributing this aspiration to Smith's argument we differ from the understanding of Smith offered in Humberstone 1992. Humberstone writes that 'The worry is not that some would-be analysis of the concept of belief fails in virtue of employing, in disguise, that very concept; for clearly no such analysis was being offered' (64). What does not worry Humberstone does worry us, as it seems to us that Smith did intend to offer such an analysis.

duced state to be a perception that not p .² But this reasoning does not count against taking the introduced state to be a perception *with the content that* not p . Someone might have a perception with the content that the road is wet without believing it is wet, for she might take her perception to be illusory. Thus Smith's appeal to perceptions with a given propositional content nicely avoids circularity. The idea of a 'perception with the content that not p ' is not entirely intuitive, but it is fundamental to Smith's approach.

Is it true that 'a perception with the content that not p ' tends to make antecedent beliefs that p go out of existence? It would tend to do so only if the agent thought that it was accurate. If the agent thought that the perception with the content that not p was misleading, there would not be a tendency for it to drive out any antecedent belief that p . Thus, before we can say whether or not there is a tendency for a perception with the content that not p to drive out a belief that p we must know more about the agent's attitudes towards the perception. We must know whether the agent takes the perception to be veridical.

One might object that there must be at least some tendency for a belief that p to be driven out of existence by a perception with the content that not p , even if the person takes the perception not to be veridical. On a functionalist account, what other tendency would give the perception the content that not p ? We reply that there appears to be no such tendency in cases of the sort we have in mind. Consider the apparent puddles one often sees when driving on a hot dry pavement. An experienced driver has no tendency at all to cease thinking the road is dry when confronted with such an appearance.

There appears to be a variety of problematic cases that cast doubt on Smith's claim about the nature of beliefs and desires. (1) Stubborn beliefs: Fred is sure it will not rain. Nicole is sure God exists. Fred and Nicole have been confronted with what they allow to be appearances with the content that it will rain or that God does not exist. Fred sees enormous black storm clouds approaching, and agrees that it looks like it will rain. Nicole says that the problem of evil makes it appear that God does not exist. But,

² It appears that it would not be circular in this way to take the introduced state to be a 'perceiving' rather than a 'perceiving that'. An anonymous referee pointed out that although Oedipus's perceiving that his mother entered the room presupposes that he believes that she entered the room, his perceiving his mother enter the room does not. Unfortunately, Smith cannot take advantage of this characteristic of 'perceivings'. His strategy requires that the introduced state have a propositional object that can interact with the agent's mental economy. The object of Oedipus's perceiving his mother enter the room is his mother herself, or the event of her entering the room, not any proposition about her. Suppose Oedipus begins by believing his mother is not in the room. He might perceive his mother enter the room in the sense at issue even if there is no tendency at all for this to drive his belief out of existence, for he might not realize that the person he perceived enter the room is his mother.

although they acknowledge that the appearances are against them, Fred and Nicole continue to be confident of their claims. Must we say that they do not believe what they think they believe but merely wish that their claims were true? (2) Beliefs in necessary truths: Jim is sure that $2 + 2 = 4$. He will not let any appearance dent this conviction. There is no introduced state that can tend to undermine it. Must we then say that Jim does not believe that $2 + 2 = 4$?

These examples deal with beliefs that do not tend to go out of existence when Smith's account says they should. But there can be desires that do go out of existence when Smith's account says they should not. (3) The fair weather fan: Sue says she roots for the 49ers. But we have noticed the team that Sue says she roots for changes frequently, apparently in response to the recent successes or failures of the team in question. Sue says that she desires that the 49ers do well. But their not doing well tends to drive out of existence this desire. Must we then say that Sue does not really root for the 49ers and instead merely believes that they will win? Perhaps it will seem that what Sue really wants is to be rooting for whichever team happens to win. We reply that even if this is what she desires at some level, she can only satisfy this desire by actually forming a desire that a particular team win. If she roots for the 49ers, she must want the 49ers to win, even though this desire will evaporate if she begins to see they are losing.

Smith suggests that, to deal with certain problematic desires, his view needs to be made more complicated by recognizing the effect a desire might have on betting behaviour. The idea is that a desire that p will be reflected in a willingness to buy tickets in lotteries that give p as a prize with a certain probability. (See Smith 1994: 208, fn. 3.) Perhaps it will seem that Smith could make use of this idea in responding to our example of the fair weather fan and perhaps also to the previous two examples concerning belief.³

We respond that some desires do not manifest themselves in a tendency to enter lotteries any more than beliefs do. This is because some things we want cannot be won in a lottery. When Sue was rooting for the 49ers, what she wanted was that the 49ers defeat their opposition in a competitive contest. This desire could not be satisfied by awarding her the prize of the 49ers being awarded the 'win'! We often want something to come about because we, or someone else, have earned it. In such cases and many others, where what we want is an outcome produced in a certain way, it can be incoherent to suppose that our desire could be satisfied by winning a lottery. In other cases, the problem is not incoherence, but that the object of our desire is something we believe no one has the power to award as the prize of a lottery.

³ This possibility was brought to our attention by an anonymous referee.

2. *Other introduced states*

Let us now return to the problem of specifying the nature of the introduced state. There is a dilemma facing direction of fit accounts. On the one hand, if the introduced state is or entails a belief, the account cannot claim to explicate the real difference between beliefs and desires. The account would be unhelpfully circular. On the other hand, if the introduced state is instead understood to be something more like a mere seeming that not p , there seems to be a wide range of cases in which the account gives incorrect results. Most obviously, the agent might dismiss the mere seeming as not veridical. In such cases, there would not be a tendency for it to drive out the agent's background belief.

This suggests that what was wanted cannot be had. The introduced state must not be a belief, but it also must not be a mere appearance or seeming. But consider yet another proposal. Perhaps the introduced state should be construed as a perception that counts as evidence that not p .

This proposal must be understood to imply that the agent herself takes the perception to count as evidence for not p . There would then be some tendency for the perception to drive out a belief that p . Of course, an agent need not believe not p as a result of what she allows to be evidence for not p . Sometimes evidence can be overturned or inconclusive. But if she allows that there is evidence for not p , she must believe that it counts in favor of believing not p . That is, the judgment that something counts as evidence for not p just is a belief. But positing this belief state to explain the distinction between beliefs and desires is again to move in a circle. It presupposes the distinction rather than establishing it.

These problems should not surprise us. Indeed, it might now seem unsurprising that we cannot find an introduced state that counts as in some way a perception with the content that not p , that is *not* itself a belief, but that interacts with the belief that p *exactly as if* it were an incompatible belief. We need such a state in order to fix without circularity whether a background state is a belief or a desire. If we are forced to concede that no suitable introduced state can be found to pull off this trick, direction of fit arguments will be unable to underwrite the distinction between beliefs and desires.

Consider now a final proposal. Perhaps we can leave the nature of the introduced state quite open in order to finesse the issue whether it is a belief or a desire. The idea is that the introduced state should be specified as merely *some* state with the content that not p . Of course it will in fact be either a belief or a desire or some other kind of state, and it will either tend to drive the background state out of existence or not. So suppose, first, that the background state is a belief that p . Then let us grant that it will tend to go out of existence if the introduced state is a belief that not p . But it pre-

sumably will not tend to go out of existence if the introduced state is a desire that not p . The more difficult case is the one in which the background state is a desire that p . It is arguable that a desire will not tend to go out of existence no matter what the nature of the introduced state. For it clearly will not tend to go out of existence upon the introduction of a belief that not p . And it might not tend to go out of existence if the introduced state is a desire that not p . For people often have conflicting desires, and are aware that they do, without either of the desires tending to eliminate the other. If this were the full story about possible background states, then when a background state that p is driven out of existence by an introduced state with the content that not p , the background state must be a belief. For, of the possible pairs of introduced state and background state that we have so far considered, the only case in which the background state would tend to be driven out of existence is one in which the background state is a belief. So understood, the direction of fit test might work.

Unfortunately, there are psychological states with propositional content that at least arguably are neither beliefs nor desires. For example, one can intend that p , hope that p , be uncertain whether p , fear that p . Suppose that the background state is an intention that p . If the introduced state were a new desire that not p , this would have some tendency to drive out of existence the prior intention that p . The introduction of a desire not to visit Paris would have some tendency to drive out of existence the intention to visit Paris. We therefore have to give up the conclusion we reached before, which was that, when a background state that p tends to be driven out of existence by an introduced state with the content that not p , the background state must be a belief. The background state could instead be an intention that p . And suppose that the introduced state is a settled acceptance of the fact that not p , or a state of contentment. This would have at least some tendency to drive out the desire that p . The desire that I see Paris would tend to be driven out of existence by the introduction of a settled contentment with my not having seen Paris. So we cannot conclude that when a background state that p is driven out of existence by an introduced state with the content that not p , the background state is not a desire. It could be a desire if the introduced state is a state of contentment. It appears, then, that the strategy we are currently considering, the idea of leaving open the nature of the introduced state except to specify its content, cannot work.

We have now considered three understandings of the introduced state, and we have argued that none of them is adequate. First, the introduced state could be some state or other with the content that not p . Second, it could be a mere seeming that not p , or something else short of a belief that not p . Third, it could be a belief. In this last case, of course, the direction

of fit account will be circular. It will depend on the application of the very concepts it is intended to explicate.

3. *Normative accounts*

One might think that a normative account of the metaphor of direction of fit could solve the problems we have been discussing, since to this point we have only considered descriptive accounts. Nick Zangwill has proposed that ‘folk psychology is normative’ and that belief and desire each have ‘a rational essence’. He holds that the difference between belief and desire is essentially normative (Zangwill 1998: 191, 193).⁴

We agree, of course, that belief and desire are *subject to* different norms. The rationality of belief is assessed relative to epistemic norms, while the rationality of desire is assessed relative to norms of practical reason. The issue raised by Zangwill, however, is the metaphysical issue whether belief and desire are *constituted* at least in part by certain norms. We will mention two problems that, it seems to us, must haunt any normative account of direction of fit. Any normative account proposes that belief and desire are constituted by differences in their normative relations either to the world or to other psychological states. The nature of the first problem depends on which of these alternatives is taken up.

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). The problem of explaining the difference between belief and desire is precisely to explain the difference between those states with propositional content that are capable of truth and those that are not. 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.

According to normative approaches of the second kind, the difference between belief and desire is a matter of differences in their normative relations to other psychological states. Zangwill proposes that a psychological state with the content that p qualifies as a belief that p only if it would be rationalized by ‘the perceptual experience that p ’, and it qualifies as a desire only if it would ‘by itself, rationalize the intention to bring it about that p ’ (195).

⁴ We are grateful to an anonymous referee for referring us to this essay.

It should be obvious that strategies of this kind must specify the nature of the ‘introduced state’ to which beliefs and desires, respectively, are normatively related. The same problems that emerged above in specifying an introduced state with the right causal role reappear in attempting to specify an introduced state with the right normative role. Indeed, our problems with Smith’s ‘perception with the content that not p ’ carry over to Zangwill’s ‘perceptual experience that p ’. Zangwill will have to specify an introduced state that is not a belief yet that normatively interacts with other beliefs exactly as if it were a belief. For reasons we have given, we doubt that there is any such kind of state.

The second problem with normative approaches is equally fundamental. We think the fact that a state is a belief or a desire *explains* why it has the normative characteristics it has, and it is not *constituted* by its having the normative characteristics it has. Moreover, a difference in the relations that belief and desire ought to have, either to the world or to other psychological states, must be due to some non-normative difference in the nature of beliefs and desires. For instance, the idea that beliefs should track the world while it is not the case that desires should track the world presupposes that, unlike desires, beliefs *can* track the world or *do* track the world. Similarly, the idea that murders are essentially at least prima facie wrong presupposes that there is some relevant non-moral difference between murders and the killings that are not murders. Murders are wrong because of some non-moral facts about the killings that count as murder. Similarly, the idea that beliefs, unlike desires, are rationalized by perceptual experiences with relevant content presupposes that there is some non-normative difference between beliefs and desires. For instance, one might say to have a belief is to take a content to be true, and perceptual experiences can be evidence of the truth of their contents.

4. Conclusion

For these reasons we think that the more fundamental difference between belief and desire is descriptive, not normative. Normative direction of fit accounts are not promising. And descriptive accounts face the serious difficulties we considered earlier. Direction of fit may be no more than a metaphor that can give an intuitive sense of the difference between belief and desire.

We have not argued against all functionalist accounts of the nature of belief and desire. Functionalism holds that beliefs and desires are states with different functional roles. Functionalism might be correct even if beliefs and desires are not cleanly distinguished by a difference in direction of fit. We think that a plausible functionalism would view beliefs and desires each as constituted by a variety of functional roles and as distin-

guished by broad differences between their roles, not merely by a difference in direction of fit. Beliefs and desires are plausibly viewed as having different inferential roles, for example, or roles in leading to changes in overall state of mind.⁵ In short, we think that neither belief nor desire can successfully be explicated in terms of direction of fit. If there is a successful functionalist account of belief and desire in this neighbourhood, it will have to be more ramified and subtle.⁶

Bowling Green State University
Bowling Green, OH 43403-0222, USA
sobel@bgnnet.bgsu.edu
dicopp@bgnnet.bgsu.edu

References

- Anscombe, E. 1957. *Intention*. Oxford: Blackwell.
- Harman, G. 1986. *Change in View: Principles of Reasoning*. Cambridge, Mass.: MIT Press/Bradford Books.
- Humberstone, I. 1992. Direction of fit. *Mind* 101: 59–83.
- Hume, D. [1739–40] 1978. *A Treatise of Human Nature*, ed. L. A. Selby-Bigge and P. Nidditch. Oxford: Oxford University Press.
- Platts, M. 1979. *Ways of Meaning*. London: Routledge and Kegan Paul.
- Smith, M. 1987. The Humean theory of motivation. *Mind* 96: 31–61.
- Smith, M. 1994. *The Moral Problem*. Oxford: Blackwell.
- Zangwill, N. 1998. Direction of fit and normative functionalism. *Philosophical Studies* 91: 173–203.

⁵ Smith acknowledges this. See Smith 1994: 115, fn. 8 (at 209–210). See also Harman 1986.

⁶ Our thanks to Janice Dowell for helpful comments on this paper.