

UNNATURAL ACCESS

BY AARON Z. ZIMMERMAN

Jordi Fernández has recently offered an interesting account of introspective justification according to which the very states that (subjectively) justify one's first-order belief that p justify one's second-order belief that one believes that p . I provide two objections to Fernández's account.

Some facts are so obvious that it is difficult to say exactly what justifies our believing them. Basic truths of logic and mathematics have this status, and so do facts about what we believe. I believe that my car is black on the basis of perception. But what justifies me in believing that I believe that my car is black? Jordi Fernández has an answer: the very perceptual state which justifies my belief that my car is black also justifies my belief that I believe that my car is black.¹ He argues for similar accounts of what justifies our second-order introspective beliefs in those first-order beliefs of ours that are grounded in memory, testimony and deductive inference. When an apparent memory justifies my belief that p , it also justifies my belief that I believe that p , and so on. In the case of perception, Fernández argues that perceptual experiences can justify both our perceptual beliefs and our second-order introspective beliefs, because these experiences are reliably connected with both perceivable facts and perceptual beliefs. When it visually appears to me that p , it typically is also the case that p ; when it visually appears to me that p , I typically believe that p . So if I believe that p when it appears that p , my perceptual belief will be most often true; and if I believe that I believe that p when it appears that p , my introspective belief will also be most often true.²

The principal advantage of Fernández's account is that it explains the so-called 'transparency' of belief: as Wittgenstein remarked, we seemingly answer questions about what we believe by turning our attention to what is true. Gareth Evans pointed out that one answers the question 'Do you believe there will be a third world war?' not by thinking about one's psychology, but by considering the current political climate.³ Fernández's theory explains this focus of attention. It is because my appreciation of global tensions both justifies my belief that war is imminent, and

¹ 'Privileged Access Naturalized', *The Philosophical Quarterly*, 53 (2003), pp. 352–72.

² Of course, one does not always believe that p when it appears to one that p – e.g., one might know that one is hallucinating – but Fernández (p. 369) thinks his account can handle this. I shall not directly discuss this complication, because my criticisms do not involve it.

³ Evans, *The Varieties of Reference* (Oxford UP, 1982), p. 225.

also justifies my belief that I have this belief, that I consult the evidence for war both in order to answer questions about what is going to happen and in order to answer questions about what I believe is going to happen.

I shall not challenge Fernández's claim that his account does justice to the felt transparency of belief, though I think transparency can be accounted for in other ways.⁴ But I shall argue that despite their appeal, the conditions for introspective justification that Fernández describes are not necessary for introspective justification, and that they could only be sufficient at the cost of being redundant.

First, I shall argue that either Fernández's conditions fail to be sufficient for introspective justification or the justifications which they supply are otiose. Suppose Mary can be in states that would justify her in believing something if she were to form that belief, but that she can still fail to believe it in such a case. For instance, we might suppose that Mary is given excellent evidence that the biological differences between species can be explained by natural selection, but that despite this evidence she remains unconvinced. Clearly, she would not be justified in believing that she believes in evolution given that she does not believe in evolution, even though she has evidence that would justify her in forming the first-order belief were she to do this. So the mere existence of evidential states which would justify one's first-order belief that *p* would not justify one's *false* second-order introspective belief that one believes that *p*. This view of the case remains unshaken even if we imagine that Mary's belief that she believes in evolution would be counterfactually dependent on the available evidence, so long as we are forced to imagine that this evidence would not lead her to believe in evolution.

In fact, the frame of mind we are here being asked to imagine would be exceedingly odd. According to Fernández's theory, Mary has excellent grounds for believing that she believes in evolution, so there is nothing irrational in her being firmly convinced (or psychologically certain) that she believes in evolution. But she does not believe in evolution. So if her beliefs enjoy normal relations with assertion, she will sincerely report that she is certain that the theory of evolution is true, but when asked whether the theory of evolution is true, she will refuse to give an affirmative answer. While this is not a genuine example of Moore's paradox (Mary will not assert '*p*, but I do not believe that *p*') the case is sufficiently like Moore's paradox to place Mary's rationality in question. Mistakenly judging that one believes that evolution is true *just because* one has excellent evidence that evolution is true is not at all like mistakenly judging of what is in fact a fake apple that it is real. Though false, the perceptual judgement is often fully justified; the false introspective judgement is of questionable rationality.

Suppose then, on the other hand, that one cannot be in states that would justify one's belief that *p* without therein believing that *p*. Then every introspective belief that meets Fernández's conditions for introspective justification will be not only justified but also true. Moreover, Fernández's account (pp. 357–8) of what it is for a belief *b* to be formed or maintained 'on the basis of' another psychological state *s* requires only (1) counterfactual dependence of *b* on *s*; (2) a disposition to appeal to *s*

⁴ See my *Directly in Mind: an Account of First-Person Access* (Ph.D. dissertation, Cornell University, 2002).

to defend *b*; and (3) sensitivity to potential reasons for thinking that the existence of *s* is not a good indication of the truth of *b*.

Given this account of the 'basing' relation, on the supposition that our justified introspective beliefs are always true, it follows that the introspective belief that one believes that *p* will always be based on one's belief that *p*. (Or at least the introspective beliefs of those of us who satisfy condition (2), in defending claims of the form 'I believe that *p*' by simply insisting that we have those beliefs we claim to have, will be based on the very first-order beliefs that make them true.) But then the justification which Fernández's account attributes to our introspective beliefs will be redundant. Why do I need to ground my second-order introspective beliefs in the states that justify my first-order beliefs, if my second-order beliefs are based on these first-order beliefs themselves?

Which horn of this dilemma does Fernández choose? He claims that there is a type of justification, namely *subjective* justification, which our beliefs must have. Does the fact that one has subjective justification for believing that *p* entail that one believes that *p*? If it does, then though one's belief that one believes that *p* might be 'based on' whatever it is that supplies first-order subjective justification, one's second-order belief will also be based on the very first-order belief that makes it true. It will then be misleading (at best) to say that the subjective justification for the belief that *p* justifies the belief that one believes that *p*. We might compare this with a case in which I know on the basis of introspection that I have a sharp pain in my foot, and I figure out its source by looking down to discover that I am standing barefoot on a broken goblet. I could cite the existence of the sharp glass to convince someone else of the truth of my introspective belief, and the glass indirectly causes my introspective belief by causing my pain, but this does not show that my belief that I am in pain is grounded in visual perception, still less that it is grounded in perception rather than introspection.

Suppose, instead, that the fact that one has subjective justification for believing that *p* does not entail that one believes that *p*. Can the existence of subjective justification for believing that *p* then justify one's *false* second-order belief that one believes that *p*? To answer this question we need to know something about what subjective justification is supposed to be. Fernández says that one is objectively justified in holding a belief when one's doing so 'maximizes truth and minimizes falsity in [one's] total body of beliefs', and that one is subjectively justified in holding a belief when one 'believes that [one] is objectively justified in holding it' (p. 368; one might question whether objective, i.e., first-order, justification should be so closely linked to truth, but this will not concern me here).

So Fernández might claim that even when one does not believe that *p*, one can still be justified in believing that one believes that *p*, but only so long as one believes that one's holding the belief that *p* maximizes truth in one's total body of beliefs. But this is curious. Perhaps there can be false introspectively justified second-order beliefs (though this is far from obvious). But what justifies one's false second-order introspective belief is surely not another second-order belief to the effect that the (in fact non-existent) first-order belief that one believes oneself to have maximizes truth and minimizes falsehood in one's overall system of beliefs. To argue otherwise is to

say that one's belief that some first-order belief exists (in one's mind) is grounded in the belief that this first-order belief has some specific non-universal contingent property (i.e., objective justification); and this gets things backwards. One believes that one's belief that p is objectively justified *because* one believes that one believes that p and that this first-order belief has a good epistemological pedigree; one does not believe that one has the first-order belief because one believes that it has the right pedigree. At any rate, this is the obvious stance to take, and Fernández has not shown it to be mistaken.

Fernández (p. 369) moves to the claim that subjective justification grounds our introspective beliefs in order to deal with examples which counter the claim that his view provides necessary conditions for introspective justification. But the move to subjective justification leaves a number of counter-examples unrefuted. Here is one. Suppose Mary's parents raise her to believe that members of a certain ethnic group E all have some negative trait N. Suppose too that when she gets older Mary has no first-order justification for believing that all Es are N. (We can imagine that when her prejudice is challenged, she lies, and claims to have repeatedly experienced Es as N when in fact she has never had any such encounters.) If she considers whether she believes that all Es are N, she will judge that she does, and her belief may well constitute paradigmatic introspective knowledge. Still, it might be that if she were seriously to consider whether or not all Es are N, and so consider whether or not her having that belief maximizes truth in her overall set of beliefs, she would lose her prejudice. For it might be that it remains in place precisely because family allegiances and habits of thought prevent her from considering the evidence. So she need not take herself to be justified in believing that all Es are N, even in the 'dispositional sense' according to which she would judge that this belief maximizes truth in her belief set were she to consider the matter. Still, she believes that all Es are N, and she introspectively knows (and so is justified in believing) that she holds this belief.

I conclude, then, that though Fernández's account resonates with the phenomenological transparency of introspective access, it cannot be right. The truth is, I think, that introspection is even more direct and unmediated than Fernández describes it to be. At least, that is what I take myself to think.

University of California at Santa Barbara