Belief and Self-consciousness

David Hunter

Abstract

This paper is about what is distinctive about first-person beliefs. I discuss several sets of puzzling cases of first-person belief. The first focus on the relation between belief and action, while the second focus on the relation of belief to subjectivity. I argue that in the absence of an explanation of the dispositional difference, individuating such beliefs more finely than truth conditions merely marks the difference. I argue that the puzzles reveal a difference in the ways that I am disposed to revise my beliefs about myself. This point develops the insight that Anscombe and others had that those of an agent’s beliefs about himself that manifest that special self-consciousness are not based on observation, testimony or inference. The puzzles show that this kind of self-consciousness involves, not a special kind of belief or even a special kind of self-reference, but a special kind of belief revision policy.

Keywords: belief; subjectivity; Anscombe; Lewis

Philosophers have traditionally sorted puzzles involving belief into two kinds. Some are treated as puzzles about what it is to believe something, about what makes a state one of belief, while others are treated as puzzles about what there is to believe, about the range of possible belief. For example, puzzles about degrees of belief, about weakness of the will, about how belief differs from desire, intention and perception, and about the causal efficacy of belief are treated as puzzles about what kind of state a belief is, and proposed solutions to those puzzles typically involve claims about the metaphysics of belief states. On the other hand, puzzles about identity, about the necessary a posteriori, about logical omniscience, and about de se and de re beliefs are treated as puzzles not about what belief is but about what there is to believe, and proposed solutions to them typically involve claims about how to distinguish logically equivalent beliefs, about how to individuate beliefs more finely than by their truth conditions alone. On my view, this treatment of the puzzles in this second group does more harm than good. Proposals to individuate belief more finely than truth conditions alone often lack independent motivation and can prevent us from seeing...
that the puzzles in this second set are really about belief itself and not about the range of possible belief.

I think that this is especially true of puzzles about first-person belief, beliefs that seem to express or manifest a special kind of self-consciousness. The puzzles seem to suggest that there can be a difference between my believing that D.H. is P and my believing that I myself am P, a difference that manifests or is grounded in a difference in self-consciousness and that makes a difference to practical rationality and autonomy. Part of what makes these cases puzzling is that there is no truth-conditional difference in the beliefs at issue. Since D.H. is P just in case I myself am P, there is no difference in what I believe to be the case whether I believe one thing or the other. But if there is no difference in how the world would have to be for each of those beliefs to be true, then how exactly do those beliefs differ? Solving the puzzle involves identifying the difference.

Plausibly, a difference in belief requires a dispositional difference, a difference in how an agent would act or react in various situations. In other work, I have developed and defended a dispositional conception of belief, but for present purposes I will simply take it for granted that if there is a difference between my believing that I myself am P and my believing that D.H. is P, then there must be a dispositional difference between believing the one thing and believing the other. It must be that my believing one rather than the other would sometimes make a difference to how I would act, react or change in the face of changing situations. Denying this would amount to saying that what is special about the first-person kinds of beliefs – and so what is special about the kind of self-consciousness that is manifested or expressed in them – makes no difference to how an agent would act or react in various situations. This sounds to me very implausible.

Given this assumption, it follows that an account of the difference between those beliefs that does not also identify a dispositional difference between them is at least incomplete. In particular, an account that distinguished the beliefs merely by assigning them different but truth-conditionally equivalent contents would be incomplete. In the absence of an explanation of the dispositional difference, such an account merely marks the difference. I think that we can understand the kind of self-consciousness that is manifested in first-person beliefs without individuating these beliefs more finely than by truth conditions. The puzzles reveal a difference in the ways that I am disposed to revise my beliefs about myself. This point develops the insight that Anscombe and others had that those of an agent’s beliefs about himself that manifest that special self-consciousness are not based on observation, testimony or inference. What the puzzles show, I think, is that the kind of self-consciousness they manifest involves, not a special kind of belief or even a special kind of self-reference, but a special kind of belief revision policy.
The puzzles that I will discuss involve an agent’s beliefs about himself. But as is well known, not all of an agent’s beliefs about himself manifest or express the kind of self-consciousness at issue. My belief that I am typing these words (or that I am staring at a computer screen) does manifest it, while my belief that I was born in London (or that I wear size 10 sneakers) does not, even though all of these beliefs are about me, at least in the sense that whether they are true depends essentially on what is the case concerning me. It seems that my belief that I am now staring at the computer screen seems to manifest or to be grounded in a different kind of awareness of myself, hence the sense that it involves a kind of self-consciousness that is not manifested or present in the case of my belief that I wear size 10 sneakers. Anscombe suggested that the difference is that in the first kind of case, but not the second, my belief need not be based on observation, inference or testimony. She even suggested that this belief can amount to knowledge even though it is not based on any evidence from these sources. But this negative characterization leaves it unclear what is special about the beliefs that do involve self-consciousness, and they leave it unclear why only those beliefs are special in that way.

It is tempting, though in the end misleading, to say that the difference is that in the one kind of case my belief is ‘from the inside’ whereas in the other case it is ‘from the outside’. This metaphor of an inside and an outside is misleading since my belief that I am typing this sentence is a belief about an event in the external world, not a belief about an internal mental event, even though (if Anscombe is right) I am able to know it without reliance on observation, testimony or inference. Likewise, my belief that I get irritated by long waits is about my inside, at least in one sense of ‘inside’, even though (it seems to me) I learned this fact about myself in the way that others do: by carefully observing my behaviour. Believing something from the inside is not at all the same as believing something about the inside. The inside–outside metaphor, though tempting, is in the end not very illuminating.

The difference between these kinds of beliefs about oneself is also not to be found in our use of the first person pronoun or of the reflexive phrase ‘I, myself’. The use of that phrase and of the first-person pronoun is neither necessary nor sufficient for the expression of a belief that manifests the relevant kind of self-consciousness. For I can just as easily say ‘I myself was born in London’ as say ‘I myself am typing these words’, even though only the second could express a belief that manifests this self-consciousness. And I can express my belief about my current activity by saying ‘D.H. is typing these words’. The difference between these beliefs has nothing to do, it seems to me, with the linguistic resources we use to refer to ourselves, either when we speak out loud or when we think quietly to ourselves. Still, it might be that the self-consciousness at issue is necessary for being able to use the
first-person pronoun or that reflexive phrase, but even so, it would not follow that it is invariably marked only by use of the first person or some special reflexive.

Though I think that the puzzle has nothing essentially to do with the use of the first-person reflexive ‘he himself’ or ‘I myself’, it will be useful in what follows to use that formulation to mark the kind of belief that manifests or reveals the sort of self-consciousness at issue. I will thus use formulations like ‘N.N. believes that he is P’ to ascribe to N.N. a belief about N.N. that does not manifest or reveal this kind of self-consciousness, and I will use formulations like ‘N.N. believes that he himself is P’ to ascribe or describe a belief that does.

The fact that not all beliefs about oneself involve the relevant kind of self-consciousness means that not all puzzles that involve an agent’s beliefs about himself are puzzles about self-consciousness. Take, for instance, John Perry’s famous story of his believing, of whoever is spilling sugar in the store, that he is making a mess, while not believing that he himself is making a mess, even though as a matter of fact the sugar is spilling from his own shopping cart. It seems that Perry believes that he is making a mess but not that he himself is making a mess. I agree that this is a puzzling case, but I am not sure that it is a puzzle about self-consciousness. It seems, rather, to be a puzzle about mistaken identity, just as Frege’s case of the astronomer who seems to believe that Hesperus is not Phosphorus is a puzzle about mistaken identity and not about self-consciousness. The astronomer seems to believe that there are more things in the heavens than there in fact are, and because of this it is difficult for us to know how to say what he believes: it seems wrong for us to use our words ‘Hesperus’ and ‘Phosphorus’ to try to characterize the two planets that he thinks there are, since we use those words to name one and the same planet. Which of the two planets in the world as he takes it to be is the one we should refer to with our word ‘Hesperus’? Why couldn’t we, since we believe that ‘Hesperus’ and ‘Phosphorus’ co-refer, also use ‘Phosphorus’ to refer to that planet? Of course, if we did, then we would fail to capture the fact that he thinks that there are two planets. Puzzles about mistaken identity, it seems to me, always involve this difficulty: it is always unclear how we who do not share the mistake, are supposed to use our linguistic resources to characterize the mistake. The same is true, I think, in the case of Perry’s puzzle. After all, it is clear that in that story Perry believes that there is some other person wandering the aisles in addition to himself, and that this other person is spilling the sugar, when in fact there is no such other person. But how can we who know that there is no such person use our words to characterize Perry’s mistaken beliefs? We cannot do it the way he tries to, through a demonstrative use of ‘he’ while pointing to the sugar, since he thinks that that use refers to someone other than himself, whereas we know that it refers to him. I am not denying that all of this is puzzling, nor am I claiming that the puzzle is fundamentally just linguistic.
My claim is only that I do not think that it is fundamentally a puzzle about self-consciousness. I will have more to say about this kind of puzzle later.

I have been suggesting that at least some of the puzzles about first-person belief that involve mistaken identity are not puzzles about self-consciousness. Even if I am wrong about this, some of the puzzles about first-person belief that do seem to reveal something about self-consciousness do not essentially involve mistaken identity. I will start with a story that focuses on action, and that seems to show that making sense of agency requires ascribing special first-person beliefs. Discussing this will lead to a second story, one that focuses more squarely on the issue of self-consciousness. These stories are hard to state clearly, and sometimes, I think, it is difficult to tell whether our puzzlement shows us something about the nature of belief or something about our ways of talking about and expressing our beliefs. Untangling the elements in the stories that show us something about how we use our words to ascribe and express belief from those elements that reveal something about the nature of belief itself, and about the nature of self-consciousness in particular, is not easy.

Here is a story that can seem puzzling. Victor and Victim are walking side by side in the woods when a bear suddenly attacks Victim.3 The two are walking close to each other on a narrow path, Victor close to the stream, Victim nearer the hill. As they walk and talk they pay very close attention to their surroundings. In fact, Victor and Victim agree on everything about their surroundings: they agree about their respective locations on the path, about the direction the bear is taking, about the level of the bear’s agitation, and about anything else that might be relevant to the scene. Let us even suppose that they desire the very same things. Each wants Victor to run for help and for Victim to drop into a ball. They are in full agreement, not just about how things are, but about how they ought to be. Unsurprisingly, Victim yells that he is under attack and drops into a ball. Victor yells that Victim is under attack and runs for help.

Here is the puzzle. It is quite tempting to think that Victor and Victim have (or at least might have) the very same beliefs and desires, since according to the story, they agree on how things are and on how they ought to be. There is nothing that one of them believes or wants that the other does not. There is no difference in how each takes the world to be. But if so, then it might seem puzzling that they acted differently, and rationally so. For, at least if we consider the intra-subjective case, it looks as though a difference in action can be rational only if there is a corresponding difference in belief or desire. That is, it seems that what it would be rational for someone to do cannot change without some change in her beliefs and desires. And yet Victor and Victim acted differently and rationally so even
though, or so it seems tempting to think, they need not have had different beliefs and desires. How can this be? Is what is true intra-subjectively – that no difference in rational action is possible without a difference in belief or desire – not true inter-subjectively? Or is it the initial temptation to think that Victor and Victim need not have had different beliefs and desires that is mistaken? But if there is a relevant difference in their beliefs and desires, then just what is it? This is the puzzle.

The intra-subjective case seems clear enough. If it is rational for S to do X, then it should remain rational for S to do it, unless there is some change in the world – either in the way it is or the way that it should be – that gets registered in S’s beliefs and desires. If it is rational for me to drop into a ball, then this should remain rational for me to do unless my assessment of the facts or of my needs changes. There can be, it seems, no change in what it is rational for an agent to do without some change in her beliefs and desires. Intra-subjectively, at any rate, it seems that a change in what it is rational for an agent to do requires a change in her beliefs or desires. But in our story, what it is rational for Victim to do (drop into a ball) is not what it is rational for Victor to do (run for help). It is not just that each does something that it would be irrational for either to do what the other did instead. (If this is not obvious, then let’s just add to the story that each believes that it would be irrational, given the bear’s attack, for Victim to run for help and for Victor to drop into a ball.) But then the puzzle is why it should be rational for them to act differently when there is no difference in their assessments of the facts or of their needs.

One response to the puzzle is to hold that there is, after all, a difference in their beliefs and desires. (From now on, I’ll focus on belief only.) Though both believe that Victim is being attacked by a bear, only Victim has the first-person version of that belief, or only he believes it in the first-person way. He would express this belief by saying that he himself is under attack. The idea is that although Victor does believe that Victim is under attack, he does not believe that in a first-person kind of way. Rather, what he believes in a first-person kind of way is that he himself is not under attack. Moreover, the response continues, this difference is relevant to their different actions. For if Victim had merely believed that Victim was under attack, without also believing that he himself was under attack, then he might not have dropped into a ball. It is his having a first-person belief that explains his actions. Admittedly, this is a difference in belief that does not involve a difference in truth conditions, since Victim’s first-person belief has the same truth conditions as Victor’s belief that Victim is under attack. So, according to this response, Victor and Victim have different beliefs, and in particular different first-person beliefs, even though they don’t disagree about the facts. And this difference in their beliefs corresponds to a dispositional difference, one that explains why they acted differently. If this response is right, then there is no puzzle as to why they acted differently.
since we can point to an apparently relevant difference in their beliefs, and consequently a relevant difference in their dispositions, to explain the difference in their actions.

Here is a different response to the story, one that does not claim that Victor and Victim have different beliefs and desires and that does not rely on a difference in their dispositions: Victor and Victim acted differently because they are in different situations, not because they have different beliefs. There is nothing puzzling about things with the same dispositions behaving differently, so long as they are in relevantly different situations. One fragile vase shatters while another does not because only the first was dropped. (Of course, agents with the very same beliefs may behave differently in the very same situation if they have different desires or emotions, but let us set that possibility aside for now.) And there is, after all, good reason to think that Victor and Victim are in relevantly different situations. Victim is being attacked by a bear while Victor is not. Something is happening to Victim that is not happening to Victor, and it is a difference that would seem to be perfectly relevant to explaining the difference in their actions. Victim did just what we would expect someone to do who believes what he does and is in a situation like his. Making sense of his dropping into a ball does not require ascribing to him any special kind of belief, but merely seeing what situation he was in and what he took the facts to be. So, on this response, we can make good sense of the fact that Victim and Victor acted differently without claiming that they have different beliefs and so without having to claim that they had different dispositions.

But is it really true that there is no relevant dispositional difference between Victor and Victim? Here is one reason to think that there is such a difference. Things with the same dispositions will act in the same way in the same situations. If two cubes of salt, A and B, are equally soluble, and if A dissolves when placed in a cup of water, then it follows that B would have dissolved if it had been placed in that water instead. But it seems that Victor would not have done exactly what Victim did, had their situations been reversed. In the original situation, Victim yelled out that he, Victim was under attack. But if the situation had been reversed – that is, if the bear had attacked Victor – then Victor would not have yelled out that Victim was under attack. He would have yelled out that he, Victor, was under attack. So it is not true that had their situations been reversed, Victor would have done just what Victim did.

I agree with this, but deny that it shows that they did not have the same dispositions in the original situation. The reason is that a thing’s dispositions can depend on its situation, and so can change as its relations to things in the world change. Any conception of belief should embrace this, since beliefs are responses to situations as well as dispositions to respond to situations. Changing situations can change beliefs, and so, given that a
difference in belief requires a difference in dispositions, changing situations can change an agent’s dispositions. Moreover, we would expect Victor’s beliefs, and so his dispositions, to have been different in the reverse situation than in the original situation. If the situation had been reversed, Victor would have believed that Victor was under attack (Victim would have believed this too), since Victor is (by hypothesis) paying close attention to the scene. Moreover, this difference in what he would have believed had the situation been reversed helps to explain why Victor would not have said in the reverse situation what Victim said in the original one. Victor would say in the reverse situation just what we would expect someone with the beliefs he would then have to in that situation. So the fact that had their situations been reversed Victor would not have said what Victim said does not show that they had different dispositions in the original situation.

Here is a second, more complicated line of thought suggesting that they had different dispositions in the original situation. In the original story, Victim is disposed to drop into a ball if a bear attacks Victim, but Victor is not. Victor is disposed to run for help if a bear attacks Victim. This certainly looks like a difference in their dispositions, and like a difference that might be relevant to explaining their different actions. But matters are not so clear. Consider again our two cubes of salt, A and B, and let’s suppose that they are molecule-for-molecule identical. They are thus equally soluble. Indeed, it seems quite plausible to suppose that they do not or at least need not differ dispositionally in any respect, since they are molecule-for-molecule identical. After all, how could there be a relevant dispositional difference without a relevant molecular difference? But matters are not so clear. Consider again our two cubes of salt, A and B, and let’s suppose that they are molecule-for-molecule identical. They are thus equally soluble. Indeed, it seems quite plausible to suppose that they do not or at least need not differ dispositionally in any respect, since they are molecule-for-molecule identical. After all, how could there be a relevant dispositional difference without a relevant molecular difference? But what about the following apparent disposition: will dissolve if A is placed in water? It seems that A has this disposition, whereas B does not: B will not dissolve if A is placed in water. But if not, then they appear to differ dispositionally after all, even though they do not differ molecularly. What is more, this difference is just like the above-mentioned dispositional difference that seems to obtain between Victor and Victim. What are we to say?

One thing to notice is that this property of being disposed to dissolve if A is placed in water is not the property of being soluble. For by hypothesis B is soluble, even though it lacks this property. And it is perhaps not inconceivable that something could have this property and yet not be soluble: it would not dissolve if placed in water, but it would dissolve if A were placed in water. I can think of no concrete example, but the case does not seem impossible. Moreover, in the case of A, its being soluble entails that it has it, and A’s having it entails that it is soluble, even though A’s being soluble is not the same as A’s having that property. A strange property indeed. One might wonder whether it is a real property at all.

But, assuming that it is a real property, it does seem clear that A’s having it is not relevant in explaining why A dissolved when placed in water. We
can, it seems, make perfectly good sense of A’s dissolution without making reference to its having that property, and without ascribing to A any property that B lacks, aside from the property of having been placed in water. Our explanation of A’s dissolution in terms of its dispositions and its situation is sufficient explanation. Reference to its having this additional property is not necessary. Nor would reference to it be sufficient, even if we also said that A was placed in water. For someone who knew only that the salt cube had that property and that it was placed in water would not yet know whether it had dissolved, since she might not realize that A is the cube in question, and so might not draw the conclusion from A’s having that property that A is soluble. There is, it seems, no role for that property to play in explaining A’s dissolution. Admittedly, explanation is partly an epistemic matter, and so the fact that reference to this property is neither necessary nor sufficient for explaining A’s behaviour might simply be a fact about our explanatory interests, and not about the cube’s behaviour. But even so, I think that this property in the case of A is metaphysically odd enough and explanatorily bizarre enough that it – and so also the analogue in the case of Victim – can safely be set aside here. Whatever lessons or challenges it has for us are not limited to psychological explanation, and should not be the basis for an account of the nature of belief.

I said that we sometimes explain an agent’s action in part by reference to the agent’s situation, that we make sense of why only Victim dropped into a ball by noting that he is being attacked by a bear while Victor is not. Of course, it is also relevant that Victim knows (or at least believes) that he is being attacked by a bear. If he had not realized this, if he had thought that the bear was going for Victor, then he might not have dropped into a ball and might instead have run for help. But Victim’s belief that he is being attacked by a bear is a belief that Victor has too. It is not as if they disagree about that. Still, while my proposed response to the puzzle does not deny that an agent’s mental states are relevant to determining the rationality of her actions, it might seem to be denying that they are the only relevant factor, since it seems to hold that external factors are also relevant. And here is a line of thought that suggests that it is mistaken about this. ‘Whether an action is rational depends on the agent’s reasons for doing what she did. To the extent that she had good reason for doing what she did, or to the extent that what she did followed from her reasons, then to that extent her action is rational or reasonable. But only mental states are reasons for action. Only an agent’s beliefs, desires, emotions, plans and such things are candidate reasons for action. Facts are not reasons for action, except as they are reflected in the agent’s beliefs or desires. So agents with the very same beliefs, desires and intentions have the very same reasons for action. No difference in reasons for action is possible without some difference in mental states.’ If this line of thinking is right, then since Victor and Victim acted differently but each with good reason, it follows that they must have
had different reasons and so different beliefs or desires, even though they
agreed on all the facts.

But the relation between an agent’s beliefs and her reasons for action is
not quite so straightforward. What may be right, I think, is that whether
an action is reasonable depends on the agent’s mental states. And it may
even be right that an agent’s reasons cannot change without some change
in her mental states. But it does not follow that agents with the same
mental states also have the same reasons for action, for which action a set
of mental states makes rational or reasonable can depend on whose
mental states they are. This is obvious in cases where agents agree on the
facts but have different desires. But even where agents have the same
beliefs and desires, what those beliefs and desires make it rational for one
to do need not be what they make it rational for the other to do. What my
beliefs and desires give me reason to do need not be what they give you
reason to do, since you are not me. Because Victim is the one under
attack, his beliefs give him reason to drop into a ball, while those very
same beliefs give Victor reason to run for help. In this way, reasons are
relative to agents in a way that beliefs are not. An agent’s reasons for
action are hers in a way that her beliefs and desires are not. She can share
her beliefs and desires without thereby sharing her reasons for action.
Admitting this does not mean holding that facts can determine whether an
action is rational independently of how those facts are reflected in an
agent’s mental states. Nor does it mean denying that an agent’s reasons
are constituted by her mental states. It means only holding that what
reasons a given set of beliefs and desires constitute may depend on whose
beliefs and desires they are. This may be puzzling, but if it is, then it is a
puzzle about how an agent’s beliefs relate to her reasons for action. Intro-
ducing special first-person beliefs is not going to resolve or illuminate this
puzzle.

But if an agent’s reasons depend on who she is, then wouldn’t she need to
know who she is in order to know what reasons she has? What if in the
original story Victor had taken himself to be Victim, and Victim had taken
himself to be Victor? In this case, they might still have agreed on all the facts
and had all the same desires, but Victor (taking himself to be Victim) would
then have dropped into a ball instead of running away. Had he taken himself
to be Victim, instead of Victor, he would have had different reasons than
those he in fact had. This line of thinking might lead to this conclusion: in
the original story, in order to have had the reasons he did have, Victor must
have taken himself to be Victor. But his taking himself to be Victor could
not just have been his believing that Victor is Victor, since he would have
believed that had he taken himself to be Victim. It must be that in the
original story, Victor somehow identified with his set of beliefs and desires
in a special way, one that made them reasons for him to run for help. This
kind of identification looks like a special kind of belief that the response we
have been considering cannot allow, since it is not a belief about the facts, but about one’s relations to one’s beliefs about the facts.

This variation on our story is quite puzzling, not least because it is not at all clear what we are to be supposing when we suppose that Victor takes himself to be Victim and that Victim takes himself to be Victor. Victor and Victim might be mistaken about where they are standing on the trail, each thinking that he is standing where the other one is in fact standing, and so both agreeing that the bear is going for Victor, though in fact it is going for Victim. We might describe this as Victor’s taking himself to be Victim, and it would make sense in this story for Victim to run for help and for Victor to drop into a ball. They would in this story have different reasons for action than those they had in the original story. Perhaps in this story Victim would have the reasons that Victor had in the original story. But that Victor has different reasons in this story is not surprising or puzzling, since Victor’s beliefs in this story are not those he had in the original story. In this story, Victor believes that Victor is on the hill side of the path rather than on the stream side, whereas in the original story he believed that Victor was on the stream side, not the hill side. Most significantly, he believes in this new story that the bear is going for Victor. I think that we can make good sense of the fact that Victor had different reasons in this revised story than he had in the original story and of the fact that in this revised story he falls into a ball even though he is not being attacked by a bear without ascribing to him any special first-person beliefs.

This story of Victor taking himself to be Victim involves his having different beliefs about the facts than the ones he had in the original story. But couldn’t he have taken himself to be Victim while not having different beliefs about the facts? Couldn’t he have agreed with himself in the original story about all the facts and still have taken himself to be Victim? Aren’t an agent’s beliefs about the way the world is independent of who he takes himself to be in that world, so that he could without any change in his view of the world change his view of who he is in that world? That this is possible may seem especially clear from a first-person point of view, for it seems that we imagine Victor reasoning as follows: ‘Whether my actions are rational, indeed what reasons I have for doing one thing or another, depends on who I am. But my conception of the world is the very same as his, since we agree on all of the facts. Indeed, our desires and hopes are the same too. There is nothing in my conception of the world, in what is available to me from the inside, to tell me whether my beliefs are reasons to run or reasons to drop. Which I should do depends on who I am. But if I were him, then I would have all the same beliefs that I have now. So who I am makes no difference to my view of the facts but all the difference to my reasons. So I cannot decide what I am to do, unless I can tell whether I am in the one situation or the other, whether I am Victim or Victor. How can I tell which I am?’
I am not sure that this is the best way to tell the story. But it does seem clear that the puzzle is not fundamentally about reasons for action, or even about action at all; it is a puzzle about an agent’s relations to her own conception of the world. In the story, Victor is not sure which perspective on the world is his perspective, even though he in some sense knows what perspectives there are. At least, he knows that his is one of two perspectives, and even knows the contents of those perspectives, but does not know which is his, and because the contents of the perspectives are the same, it is not that he is unsure what he believes. Rather, he is unsure which believer he is. Perhaps the uncertainty is not essential to the story; perhaps it would be enough if Victor took it to be conceivable that he might have had Victim’s perspective. This would be like Thomas Nagel’s claim that in some moods he finds it remarkable that he is Tom Nagel, as if the perspective he has on the world is in some way contingent, and not just contingent in the sense that he might have had different beliefs and desires. If this kind of uncertainty is possible, then it will give rise to difficulties about rational action, since it is hard to imagine deciding how to act if one is not sure what perspective is one’s own. But it is the idea of being in this way detached from one’s own perspective, of being unsure which set of beliefs is one’s own, or of finding it surprising or remarkable that one has the perspective one does, that is at the heart of this new puzzle. To address it, I want to discuss a different story, one that can, I think, help us focus more sharply on this issue.

The story is David Lewis’ tale of the omniscient Gods, where one is on the tallest mountain and the other is on the coldest mountain. Each God knows all the facts there are to be known, including any facts about each God’s perspective on the world, and even about what it’s like to be that God, if there are such facts. But one day the God on the coldest mountain wonders to himself, ‘Which God am I? Am I the one with the tallest-mountain perspective on the world, or the one with the perspective from the coldest mountain? I know there are these two perspectives, but which perspective is my perspective? Oh, I’m that God! I am the God on the coldest mountain! That perspective on the world is my perspective.’

The puzzle is how we can make sense of the content of the imagined God’s uncertainty. He seems to be experiencing a real uncertainty, but he lacks no factual information since he is omniscient. His uncertainty is thus not a factual uncertainty. So what does he learn, or come to believe, when he discovers which perspective on the world is his perspective? It seems that he acquires a new belief, but not a belief about a fact that he didn’t already have a belief about, since by hypothesis he was omniscient and knew all the facts. Nor does he come to know what he believes, since he already knows that his is one of two perspectives on the world and what the world looks like.
like from both. What he needs to learn is which of those two perspectives, which are the same from the inside, is his.

A common response to this story is that it shows the need for self-locating beliefs, for beliefs that identify for an agent her own perspective on the world, and that embodies or expresses the kind of self-consciousness that the God acquires in our story. While he has plenty of beliefs about himself, until he acquires a self-locating belief none of his beliefs amount to the relevant kind of self-consciousness. According to this response, both Gods agree that the uncertain one is on the coldest mountain, but only he comes to believe this in the first-person, self-locating kind of way. The other God does not. The God on the coldest mountain comes to learn, as he would put it, that he himself is on the coldest mountain. When he gains this belief, his doubts are removed. This is the kind of belief that Victor in the story at the end of the last section also seems to need. In order to decide what to do, he needs to know not just that the bear is attacking Victim, which is something Victim knows too, but that the bear is not attacking himself. It looks as if, from the first-person point of view, I need to know which perspective on the world is my perspective, and it seems that this requires more than simply knowing how things look from my perspective, since the world might look the same from different perspectives. Something in my perspective on the world must mark it out as mine. I need a self-locating belief, a belief that only I have, a belief that is therefore individuated more finely than by truth conditions alone. The moral is that only an account of belief that embraced special first-person beliefs individuated more finely than by truth conditions could account for this kind of self-consciousness, for what I know when I know which perspective on the world is mine.

Ultimately, I will challenge this moral, but I want to pursue the puzzle for a bit first. That the Gods are omniscient is crucial to the story. For it is supposed to be a story about agents who are certain about all the facts, and yet still uncertain about something. If this kind of uncertainty is coherent, and if the imagined uncertainty is eliminated by the God’s acquiring some belief, then the content of that belief cannot be captured except by individuating beliefs more finely than by truth conditions alone. If, however, the story in fact involves some hidden factual uncertainty, an uncertainty about how things are, then the story will not constitute a reason to endorse special first-person beliefs at all. So it is crucial to the story that the God have no factual uncertainty at all.

I think that part of what fuels the puzzle is the distinctively philosophical idea that having a belief is like having a map of the world, an idea that can be traced back at least to Ramsey and more recently to David Armstrong. The idea is appealing in part because it seems to capture some of what we have in mind by the objectivity of belief, for beliefs seem to represent the world as if from no perspective in particular just as a map may represent the world or some part of it in such a way that anyone can use it, regardless of
who or where she is in the world the map is about. This independence secures a kind of objectivity to the map, and the idea that having a belief is like having a map seems to capture one sense in which beliefs are objective: anyone can have the same beliefs about the world. But if we think of having a belief as like having a map, then the God's imagined uncertainty can seem possible, indeed even inevitable. For one can have a map and yet not know where one is on the map, or in the world as the map represents it as being, even when the map in fact accurately portrays one’s location. A map in a shopping mall might accurately represent the location of all the stores in the mall, but this is not enough to determine one’s own location in the mall. Indeed, even if the map represented the location of all the shoppers in the mall, this would still not be enough, since one might not be able to tell from the map alone which of the represented shoppers is oneself. For this reason, maps in malls typically have a ‘You are here’ sticker, identifying the location of the person looking at the map. If we think of having a belief as like having a map of the world, then it will seem that we will need self-locating beliefs as like we need ‘You are here’ stickers on maps. A self-locating belief, one that expressed or manifested the relevant kind of self-consciousness, would identify which of the many perspectives on the world we believe there are is in fact our own. Only such a belief could remove or preclude the kind of uncertainty that we imagined the God experiencing.

There are different accounts of the nature of self-locating beliefs. Gottlob Frege based his account on the idea that all of us are presented to ourselves in thought in a special way, one in which we are not presented to anyone else and (presumably) in which no one else is presented to himself. The way in which I am presented to myself in thought, which Frege assimilated to a kind of self-reference, locates only me and does it only for myself – for this reason it is self-locating for me. The difference between Victim’s believing that Victim is under attack and his believing that he himself is under attack is that only the second belief involves his thinking of himself in that special first-person kind of way. His belief must be individuated in part in terms of that ‘way of thinking’, and so not only in terms of its truth conditions. Frege was a propositionalist, and held that the special ‘way of thinking’ forms part of the content of the object believed. But other propositionalists have taken first-person ways of thinking to be a matter of one’s relation to the object believed, so that while Victim believes the same proposition in both cases, his relation to that proposition is different when he believes that he himself is under attack. But even on this alternative view, that belief cannot adequately be characterized in terms of its truth conditions alone. On either of these views, only some beliefs are self-locating. But David Lewis goes further and holds that each of an agent’s beliefs is self-locating, that each belief has to be individuated in part by the agent whose belief it is. On this view, each one of my beliefs includes a kind of index identifying me as its believer. Each of my beliefs is thus in a sense about me and as a consequence...
no one else can share my beliefs, since no one else can share my self-locating beliefs, even if our conceptions of the world are equivalent.

These proposals face difficulties of their own. I have already noted that individuating Victor and Victim’s beliefs more finely than by truth conditions conflicts with the idea that beliefs make a difference to dispositions, since there is good reason to think that Victor and Victim might well have the very same dispositions. My point is simply that it is not clear what we are saying about an agent when we ascribe to her a self-locating belief. Without some insight into what dispositional difference we are trying to identify or mark in ascribing self-locating beliefs, talk of them risks merely tagging a difference and not illuminating it. Moreover, if Frege is right, then no one else can share some of my beliefs about myself, and indeed may not even be able to entertain them. This seems hard enough to accept, without having to accept Lewis’ extreme view that none of us can ever have any beliefs in common. More importantly, though, I doubt that these proposed solutions can really help us to understand what is genuinely puzzling about the story of the Gods and about Victor and Victim.

As a first step to seeing this, we should note that the imagined God cannot really be omni-certain. For we know that he is uncertain about how his words and thoughts hook up with the world. Or rather, he is uncertain about the truth value of what he says or would say in uttering the sentence ‘I am the God on the coldest mountain’. He does not know whether in saying that he is saying something true or something false, and this uncertainty is about the facts. But he would also be uncertain about what it is that he is saying in saying that, since he is uncertain about which God the ‘I’ refers to. He knows that he is saying either something true about the one God or else something false about the other God, only he doesn’t know which thing he is saying. He is uncertain about the truth conditions of those words and thoughts, and this is an uncertainty about the semantic facts. Because he is not really factually omniscient or even omni-certain, the story constitutes no special support for the need for special first-person beliefs after all, for his is not a case where an agent is certain about all the facts yet uncertain about something. The story of the Gods would provide support for this only if the Gods were factually omniscient (or omni-certain) because in that case the imagined uncertainty could not be viewed as uncertainty about the facts. Since the God is uncertain about the facts, we can make sense of the uncertainty without individuating beliefs more finely than truth conditions. In which case the story does not constitute an argument for individuating beliefs more finely than truth conditions.

The point is not that those semantic uncertainties are the content of his uncertainty. It is not that in specifying the semantic facts that he is uncertain about we have completely characterized the God’s imagined uncertainty. I think that part of the God’s puzzlement is a confusion about identity, and one feature of such cases (as I said in section 1) is that we who do not share
the confusion cannot always adequately use our linguistic resources to say how the confused agent takes the world to be, and so what the content of his uncertainty is. There is admittedly something unsettling in the idea of a disagreement about the facts that we cannot state, but recognizing limits to belief ascription does not force us to individuate beliefs more finely than truth conditions. In any case, my present point is simply that one can see that the God’s semantic uncertainty prevents his case from providing support for special first-person beliefs without having to say what he is really uncertain about. One can accept the semantic diagnosis as a solution to the puzzle without seeing it as providing the content of the God’s beliefs.

More importantly, though, if the God were really uncertain which perspective on the world were his, then neither self-locating beliefs nor anything else could remove it, for the uncertainty would simply recur at a higher level. Suppose that the God on the coldest mountain acquired a self-locating belief, one that identified for him which perspective on the world was his perspective. Being omniscient, each God would know that each God had a self-locating belief, and indeed would know which God had which belief. But couldn’t we then imagine the God of the coldest mountain wondering to himself as follows: ‘well, I know that there are two Gods and I know that each one has a self-locating belief, but which of them locates me? Which of these two beliefs identifies my perspective on the world? Am I the God with this self-locating belief or the God with that self-locating belief? Isn’t it somewhat remarkable that one of these two self-locating beliefs locates me?’ If the original uncertainty were possible, what would prevent the occurrence of such a second-level uncertainty?

Compare this to the case of ‘You are here’ sticker on the map in the mall. The sticker cannot all by itself remove one’s uncertainty about one’s location in the mall, since one might wonder whether the ‘you’ refers to oneself. Maybe teenagers moved the location of the sticker on the map just to trick unwary shoppers. The fact that the map has a ‘You are here’ sticker does not by itself preclude uncertainty as to one’s location. Indeed, if one really knew nothing at all about one’s location, then the map would be of no help at all, no matter how many ‘You are here’ stickers it had. Those stickers help only because one already knows something about one’s location. To use a mall map, one has to know already that one is looking a map in the mall, and that the ‘you’ on the sticker is to be taken as referring to anyone reading the map.

This is an important difference between the uncertainty of the lost shopper and that of the God, one that brings into question the intelligibility of the God’s imagined uncertainty. When we think of the omniscient God, we are supposed to think of him as like the lost shopper staring at a complete map of the world, wondering who he is in the world of the map, except that unlike the lost shopper who knows that he is shopping, the God is omni-uncertain about his own location. The shopper is able to use what
he does know about his location to triangulate with what the map says to figure out where he is in the mall. His uncertainty is intelligible to us because we can conceive of him triangulating in this way. But the God is supposed to have nothing to go on but his conception of the world, and so for him there is no triangulating; his uncertainty is completely detached from any knowledge of his own location. But why should we think that the lost shopper’s intelligible uncertainty could survive being completely detached? Another way to put this is simply that, if the imagined God’s detachment were possible, what could bridge it? If uncertainty as to which perspective on the world was one’s own were really possible, then, it seems to me, nothing could remove it. This shows that the puzzle is not a special reason to introduce special first-person beliefs at all, since introducing them would not help us to understand why the uncertainty does not occur.

In a way, David Armstrong suggests as much when in discussing the idea that having beliefs is like having a map of the world he says that beliefs, unlike maps, do not need interpreting. One of the things we need to interpret in interpreting a map is our own location on it we need to figure out where the map represents us as being. We need to understand the map. But unlike maps, we do not need to interpret our beliefs, for as Armstrong puts it, beliefs are our interpretation. But if we do not need to interpret our beliefs, then why do we need self-locating beliefs? If we do not need to interpret our beliefs about the world, then why do we need to identify where they represent us as being? The idea that self-locating beliefs are needed stems wholly, it seems to me, from the idea that an agent has to find himself in the world of his beliefs, that having beliefs is like having a map of the world, a map about which we can ask the question where am I in this conception of the world? But if I do not need to interpret my beliefs, if I do not need to find myself in the world as I believe it to be, then I do not need to ask this question, and I do not need self-locating beliefs.

The uncertainty that I am claiming is not coherent is easily confused with uncertainties that are possible. I could have been somewhere else, and so I might be uncertain where I am. And I could have believed myself to be somewhere else, and so it is contingent that I believe myself to be where I do. So I can ask why am I in the location I believe myself to be in, or why it is that I believe myself to be in that location. These are like the questions Nagel asks when he notes that it is a somewhat remarkable fact that he is Tom Nagel. It is contingent that I have the perspective I do, but only in the sense in which the view of out a given window might have been different. That I have the window on the world I do is not contingent. Being uncertain whether I am where I take myself to be is not the same as being uncertain whether I am who I take myself to be. Finding it remarkable that I have the beliefs that I take myself to have is not the same as finding it remarkable that I am the believer that I take myself to be.
I have been arguing that individuating beliefs more finely than by truth conditions does not help us to understand what is puzzling in the story of the Gods. The puzzle is about how belief can constitute a subject’s perspective on the world and yet be objective at the same time. The demands on subjectivity seem to be incompatible with those on objectivity. Objectivity requires that belief be from no perspective in particular, so that anyone could have that very belief, like a map that anybody can use. But subjectivity requires that my beliefs contribute to, or partially constitute, a unique perspective on the world, one occupied by me alone, like a map that only I am in a position to use. How is it that belief can satisfy both these demands? That is the puzzle.

Some distinctions can help us to see part of the answer. We should distinguish the content of the belief from the believing of the content – what one believes from the fact that one believes it. The content is objective in that anybody can have a belief with that content. Having a belief with a given truth-conditional content does not require having some specific perspective on the world. But the believing is subjective. My believing something is not the same as your believing it. The fact that I believe something is distinct from the fact that you believe it, even if you and I believe the very same thing. Perhaps we should understand what’s subjective about belief in terms of the states of believing and not in terms of what is believed. Perhaps subjectivity is found in the subject and not the object of the belief. Here again it seems that individuating beliefs more finely is not going to help us to understand what is essentially subjective about belief.

A second distinction is more suggestive. It’s a distinction between what is conceivable by me versus what is believable by me. You and I may agree on all the facts and even about what is possible. We may agree about how things are and about how they might have been, about what possible worlds there are. In particular, we can agree that I might not have existed, that the set of possible worlds includes a D.H.-free world. Others can believe that I do not exist. They can believe that a D.H.-free world is the actual world. But I cannot believe that I inhabit such a world. The worlds I can take to be actual must all include D.H., and this limit applies only to D.H. This marks a given perspective on the world as D.H.’s perspective: it is the one from which he exists in all the worlds that are believable from that perspective. This is a limit on my perspective on the world. But notice again that we can describe this fact about my perspective on the world in terms of truth conditions alone. This difference remains, and can be identified, even if you and I agree on all the facts not just about what is actual but about what is possible. It is a difference in the limits of belief, a difference that marks one perspective as yours and another as mine. It is also a difference that makes a difference to our possible dispositions, since it marks a difference in how I can treat the world. I cannot treat the world as if it were one in which I did
not exist; anyone else could. Perhaps the essence of my perspective on the world is the way it limits what can be believed from it?

The point about the limits of belief leads to a point about the limits of belief revision. I can revise my beliefs in many ways, but I cannot revise my conception of the world in such a way that my conception no longer includes me. You can, of course, revise your conception of the world in that way. You might gain evidence that I no longer or never did exist, and might change your conception of things accordingly. But I cannot, and this difference in how we are each disposed to revise our beliefs marks our own perspectives on the world: my perspective is the one from which my existence cannot be up for revision. This fact about my perspective is enough to capture a part of what is involved in my being self-conscious. I am aware of myself, and keep track of myself, in a way that is fundamentally different from the way in which you are aware of me and keep track of me. The difference is not a difference in how your beliefs are about me, or in the way you refer to me in thought and judgment. The difference, rather, is a difference in the limits of possible belief revision. The source of self-consciousness, I am suggesting, resides in the limits of belief revision, and not in any special kind of reference or belief.

But isn’t there still room for a remaining, lingering uncertainty? For can’t we conceive of D.H. wondering to himself as follows: ‘I know that this perspective is the D.H. perspective on the world. I know that from it all the believable worlds include DH, and that from it no revision can remove D.H. from the world. But why is this perspective my perspective? Isn’t it somewhat remarkable that it is? For couldn’t I have instead occupied one from which all the believable worlds included Tom Nagel?’ If this line of thought is coherent, then doesn’t this show that there remains a kind of uncertainty that only an appeal to special first-person beliefs can make sense of?

I am inclined to doubt that this imagined uncertainty is really coherent. I think that we have been misled by the man in the mall analogy. His uncertainty is conceivable, but only because it requires his knowing something about his location in the mall. We can make sense of his being uncertain as to his location in the mall as he stares at the map precisely because we think that he knows something about his location, if only that he’s looking a map in the mall. But the man in the mall’s uncertainty is not the uncertainty that we are supposed to be imagining when we imagine an agent who does not know which perspective on the world is his own. For that is to be a completely detached uncertainty, one that does not rest on some certainty about one’s location. It is a kind of uncertainty for which there is no triangulating. This kind of detached uncertainty is, I think, probably inconceivable. But even if it were conceivable, it is not clear what it would show us about the nature of belief. For it is one thing to conceive of feeling it and another actually to feel it. I’ve never felt that uncertainty, and I doubt that anybody else has either. Feeling the uncertainty would be philosophically significant. We would then
have to have some account of the content of the uncertainty and of what kind of belief could relieve or preclude it. But merely conceiving of feeling the uncertainty is not obviously of philosophical significance. Perhaps it is but one more of those misleading feelings that we learn to live with in philosophy even though we know that it teaches us nothing. Perhaps the felt gap between the objective and the subjective is like an illusion or hallucination that we cannot avoid experiencing even though we are no longer tricked by it.¹²

Ryerson University, Toronto, Canada

Notes

3 The story is a variation by Cara Spencer of one first sketched by John Perry. What I have to say about the story, and indeed what I have to say about de se belief itself, owes a considerable amount to several discussions with Cara over the years.
4 One complication that I will simply note is that whether an agent’s beliefs and desires constitute reasons for action may depend on whether she identifies with those beliefs and desires. This is a point forcefully argued in Harry Frankfurt, The Importance of What we Care About: Philosophical Essays. (New York: Cambridge University Press, 1988). I discuss the case of identifying with one’s beliefs in ‘Belief, Alienation and Intention’, unpublished paper.
5 One might also think that practical reasoning – reasoning about what to do – requires the use of the first-person pronoun and that this use in turn requires having special first-person beliefs. I doubt both claims. But I also doubt that the kind of self-consciousness at issue requires the ability to reason about what to do. I think that many animals and most pre-linguistic humans are self-conscious in the relevant sense – they have conceptions of the world and even act intentionally in it – even though they do not engage in practical reasoning. I am willing to allow that such reasoning might typically involve the use of first-person pronouns and even of reflexive phrases like ‘I myself’. But, as I noted in section 1, the difference between beliefs that manifest the self-consciousness at issue and those that do not is not invariably marked by our use of such words or phrases. It is more primitive than our linguistic capacities. For an interesting recent account of self-consciousness that attempts to ground it in practical reasoning, see Sebastian Rodl’s Self-Consciousness, (Cambridge, MA: Harvard University Press, 2007).
BELIEF AND SELF-CONSCIOUSNESS


12 This paper has benefited from the advice and comments of many friends and colleagues. I have given versions of it as papers at the York University, the University of Guelph, Ryerson University, and at annual meeting of the Canadian Philosophical Association. I am indebted to the audiences on these occasions for their helpful questions and comments. I am also indebted to anonymous referees for this journal. My thinking on these topics is heavily indebted to discussions over many years with Cara Spencer and, more recently, to discussions with Gurpreet Rattan, Robin Jeshion and Sergio Tenenbaum.