

# Belief, Information and Reasoning

Bruno Whittle

Here are two plausible ideas about belief. First: beliefs are our means of storing information about the world. Second: if we believe something, then we are willing to use it in reasoning. But in this paper I will introduce a puzzle that will show that these cannot both be right. The solution, I will argue, is a new picture of belief, on which there are two different kinds of belief—one for each idea.

The puzzle is based on a certain sort of example that is familiar from discussions of knowledge: in particular, the main example will involve someone first thinking about what they are going to do next summer, and then about whether to buy a lottery ticket. But I will argue that we have been thinking about these examples in the wrong way: rather than focussing exclusively on knowledge, we should *also* have been asking what they tell us about belief. Only once we do that, I will argue, will we be able to see what is really going on in these examples.

The structure of the paper is as follows. In §1 I will introduce the puzzle. In §2 I will outline the new picture of belief, and explain how it solves the puzzle. But in §3 I will consider whether one might not be able to solve the puzzle in a less radical way (i.e., without giving a new picture of belief). I will argue, however, that no such solution would seem to be available. In §4 I will then return to the proposed picture, and try to go some way towards filling in, and refining, the outline that I initially gave. Finally, in §5, I will try to give an alternative argument for the new picture, using considerations from desire.

## **1. The Puzzle**

The puzzle that I will raise, then, stems from two apparently plausible ideas about belief. The first of these is:

(I) Beliefs are our means of storing information.

Thus—the idea is—the way in which we store information about what the world is like is by having beliefs about it. For example, the way in which we store information about what snow is like is by having beliefs about it (such as the belief that snow is white).

The second idea is: if we believe something, then we are willing to use it in reasoning. I.e.:

(R) If S believes that p then she is willing to use p in reasoning.

For example, suppose that Rachel wants to build a snowman, and that the ground to her left is green, while that to her right is white. And suppose also that she believes that snow is white: then (the idea is) she will use this proposition in her reasoning (and as a result will head right).<sup>1</sup>

(I) and (R), then, would seem to be very plausible ideas about belief. We will see, however, that they cannot both be true. To see why, suppose that they are, and consider the following example.<sup>2</sup> To begin with, suppose that Rachel is thinking about what she is going to do next summer. And suppose that what she really wants to do is go on safari. *But*, after checking her bank balance, and thinking about how much she is going to earn between now and then, she reluctantly concludes that she is going to have to stay in the US. That is, she comes to believe that she will be in the US next summer.

But suppose next that she goes to a deli to buy a sandwich. And suppose that while she is there she thinks about buying a lottery ticket. She thinks: if I win, then I will be on safari next summer; and she decides to buy the ticket. Now, in this new situation, Rachel is *not* willing to use the proposition that she will be in the US next summer in her reasoning: because if she was, she would conclude that there is no point in buying the ticket (for she would reason that if she buys the ticket and it wins, then she will be on

---

<sup>1</sup> (R) would ultimately have to be qualified to take account of (what one might call) ‘overriding factors’. For example, suppose that S is very lazy, or that she believes that reasoning is evil: in such cases, she may believe something, but be unwilling to use it in reasoning. Thus (R) would have to be qualified so as to make a claim only about what happens in the absence of such factors. But I will not try to state this qualification precisely here, since the issue is orthogonal to the concerns of the paper (and, further, any account of what we are willing to use in reasoning would seem to face a similar problem).

<sup>2</sup> For discussions of this sort of example in connection with knowledge see, e.g., Cohen 1988, Harman 1973, Hawthorne 2004 and Vogel 1990.

safari, and not in the US, next summer; thus, since she *will* be in the US, if she buys the ticket it won't win). Thus it follows from (R) that she no longer believes that she will be in the US next summer. But then it follows from (I) that she has lost the information she acquired at the start of the example: for if our store of information is simply our beliefs, and Rachel has lost the belief she acquired, then it follows that she has lost the information she acquired.

But of course that is not right. For suppose that when Rachel gets home she takes a call from a friend who wants to know if she is going to be around next summer. Rachel will be able to answer 'Yes' immediately. But that means that she still has the information she acquired at the start of the example: because it is still available for her to use to guide her answer; after all, it is not as if she has to revisit the facts about her bank balance and so on; rather, she simply uses the result of her earlier deliberations. Thus, as promised, (I) and (R) cannot both be true.

And this is far from an isolated counterexample: rather, others are easy to come by. For one such, suppose that David is thinking about when to buy a new computer. And suppose that he decides to keep his current machine until next year. That is, he comes to believe that he will keep his computer until next year. But suppose now that he moves on to thinking about insurance. He will not, in this new situation, be willing to use in his reasoning the proposition that he will keep his computer until next year: for example, he will not be willing to reason that since he is going to keep it until next year, it is not going to be stolen before then, and so there is no need to insure it. Thus, as before, it follows from (R) and (I) that he has now lost the information he acquired at the start of the example. But, again, this isn't right: for if, after he has finished thinking about insurance, a colleague calls and asks if he is going to have his computer later in the year, he will be able to answer 'Yes' immediately; so it follows that he has *not* lost the information after all. Thus, again it follows that (I) and (R) cannot both be true. And many other examples, which can be used to make the same point, are easy to find.

## 2. A New Picture

The solution, I want to argue, is a new picture of belief. On this picture, there are two distinct components: on the one hand, there is our information; and, on the other, there are the things that we are taking seriously at any given point. Each of these can be thought of as a set of worlds.<sup>3</sup> Thus, on the one hand, we have an ‘information set’: consisting, intuitively, of those worlds that our information leaves in play. But, on the other hand, the idea is that at any given point we are only taking *some* worlds seriously; that is, we are only taking some worlds seriously as candidates for how the world is. For example, in a typical situation in which Rachel is thinking about what she is going to do this weekend, she might take seriously (typical) worlds where she goes to visit her parents, and (typical) worlds where she stays at home; but she will *not* take seriously worlds at which she is hit by a bus on Friday morning, or at which there is a nuclear holocaust on Friday afternoon. Thus, in addition to an information set, we also have a ‘relevance set’: consisting of those worlds that we are (at that point) taking seriously. Thus, the worlds that are not in S’s relevance set are those that she is not taking seriously, but it will be useful to have a similar term for the worlds that are not in her information set; so I will say that S has ‘ruled out’ a world if it is not in this latter set.

These two sets then give rise to two distinct kinds of belief, as follows. For example, consider again Rachel’s deliberations about what she is going to do this weekend. And suppose that Rachel decides to visit her parents. What is going on in this sort of case—according to the picture I am proposing—is that Rachel is ruling out (i.e., eliminating from her information set) the *relevant* worlds at which she does not visit her parents this weekend. Thus, the only relevant worlds *left* are those at which she does visit them. However, it is not that she rules out absolutely *every* world at which she does not visit them (for example, she does not rule out the far-fetched world at which she is hit by a bus on Friday morning). Thus, in one sense, Rachel believes that she will visit her

---

<sup>3</sup> In this paper, by a ‘world’ I will mean a metaphysically possible world, i.e., a metaphysically possible way the world might have been. And I will state the proposed picture of belief in terms of these. Ultimately, one might very well want to replace this notion of a world with a more liberal one. However, I will not, in this paper, consider how exactly one might do this. Rather, I will stick with the notion of a metaphysically possible world because it is both familiar and sufficient for my purposes here. However, it should be clear that there is nothing about the general approach that I am proposing that requires using this as opposed to a more liberal notion.

parents this weekend: i.e., in the sense that this proposition is true at every *relevant* world that is left in her information set. However, in another sense, she does *not* believe this: because there are still *some* worlds in this set where this is not true (such as the hit-by-a-bus-on-Friday-morning world). Thus, on this picture, there are these two distinct kinds of belief. On the one hand, there are *R-beliefs* that are determined by our information together with what we are taking seriously; that is, we R-believe something if it is true at every *relevant* world in our information set; thus, in the example, Rachel *does* R-believe that she will visit her parents this weekend. But, on the other hand, there are also *I-beliefs* that are determined by our information alone; that is, we I-believe something if it is true at *every* world in our information set; so, in the example, Rachel does *not* I-believe that she will visit her parents.<sup>4</sup>

The idea is then that (I) is true of our I-beliefs (i.e., it is *these* that are our means of storing information) while (R) is true of our R-beliefs (it is *these* that are connected to reasoning). Thus:

(I\*) I-beliefs are our means of storing information.

(R\*) If S R-believes that p then she is willing to use p in reasoning.

Thus, in the example that we have been considering, it is *not* part of Rachel's information that she will visit her parents this weekend; nevertheless—despite this—this proposition *is* something that she is willing to use in reasoning (for example, if a friend calls and asks what she is going to be doing this weekend, she will use it in her reasoning about how to answer). And the idea is that, quite generally, this is how things go: we store information by having I-beliefs; but we are nevertheless willing to use in our reasoning anything that we merely R-believe.<sup>5</sup>

---

<sup>4</sup> This picture of belief is structurally similar to the account of *knowledge* proposed in Lewis 1996 (on that account, 'S knows that p' is true in context C iff S has eliminated all of the not-p-worlds that are relevant in C). Thus one might think of the present paper as arguing that this basic structure goes deeper than Lewis realized: applying not merely to knowledge, but to the apparently more fundamental state of belief itself. (However, I should note that a further difference between the two proposals is that, on Lewis's account, knowledge-ascriptions are *speaker*-relative; whereas, on the proposal of this paper, the corresponding notion of belief (R-belief) is *subject*-relative; but on this see §4.3.)

<sup>5</sup> Note that R-believing is entailed by I-believing (because if something is true at absolutely every world in our information set, then, a fortiori, it is true at every *relevant* world in that set).

That, then, is the picture of belief that I want to propose; in a moment I will explain how it solves the puzzle; but first I want to say something about my decision to call both I-beliefs and R-beliefs kinds of *belief*. For while it does seem to be natural to think of each of these as a variety of belief, all that I *really* want to claim here is that we have these two distinct kinds of mental states, one of which is tied to information, and the other of which is tied to reasoning; and it is then a further matter which of these should really count as beliefs (or which of these the English word ‘belief’ really applies to). Thus, one view (the one which my presentation has so far assumed) is that *both* I-beliefs and R-beliefs are beliefs. But one might also take the view that only I-beliefs are, or that only R-beliefs are (and so on). Thus, although in this paper I will, in my terminology and more generally, assume the first of these views, this is simply for the purposes of presentation, and everything that I say could very easily be rephrased so as to be compatible with any of the alternative views about this.

I will now explain how the proposed picture solves the puzzle. Thus, consider again the main example of §1, starting with the part in which Rachel is thinking about what she is going to do next summer (and checking her bank balance and so on). The idea is that, at this point, Rachel is taking seriously worlds at which she has saved enough from her salary to be able to go on safari; and she is also taking seriously worlds at which she has not saved enough, and at which she instead spends the summer in the US. But she is *not* taking seriously worlds at which she is kidnapped and forced to work as a safari guide; or worlds at which she wins the lottery and uses the winnings to go on safari. What Rachel does (in this part of the example) is then to rule out those worlds that she is taking seriously and at which she is not in the US next summer. On the other hand, she does *not* rule out any of the more far-fetched worlds that she is not taking seriously (and so these remain in her information set). Thus, at the end of this part of the example, Rachel R-believes that she will be in the US next summer (even though she does not I-believe it).

In the next part of the example—when Rachel is in the deli and considering whether to buy the lottery ticket—what happens is simply that certain worlds that *were irrelevant* become relevant: for example, worlds at which she wins the lottery and uses her winnings to go on safari next summer. Thus, since she has not ruled out these worlds, she no longer counts as R-believing that she will be in the US next summer: and so (R\*)

no longer gives the result that she is willing to use this proposition in her reasoning. But—crucially—we get this change despite the fact that her information remains the same (in particular, the worlds that she ruled out in the first part of the example *remain* ruled out). Thus, unlike when we had (I) and (R), we can now have a change in what Rachel is willing to use in reasoning *without* a subsequent change in her information (because (I\*) and (R\*) tie reasoning to R-beliefs, while information is instead tied to I-beliefs).

Consequently, on the proposed picture, the information that Rachel acquired in the first part of the example *is* still available for her to use in the third part: meaning that we no longer have a puzzle. More precisely, what happens in the third part of the example—on this picture—is simply that the worlds that became relevant in the second part go back to being *irrelevant*. As a result, Rachel goes back to R-believing that she will be in the US next summer, and so by (R\*) we get the desired result that Rachel will once again be willing to use this proposition in her reasoning. But, crucially, we get this *without* having to say that Rachel must have ‘reacquired’ the information she acquired in the first part.

Thus, on the proposed picture, we get all the right results about when Rachel acquires and loses information, and about what she is willing to use in her reasoning—while retaining versions of (I) and (R). So the puzzle would seem to be solved.

Now, as I said in the introduction, the sort of example that we have been considering has been much discussed in connection with knowledge. However, if what I have said so far is on the right track, then it would *seem* that we have been thinking about such examples in the wrong way. For, even before we get to questions about what Rachel *knows*, there is a puzzle relating purely to what she believes. And, further, it is hard to imagine that what we ultimately want to say about knowledge in such cases will not to a large extent be affected by what we say about belief. Thus, to focus principally on what these examples tell us about knowledge, as we have been, would seem to be the wrong way to going about things.<sup>6</sup>

---

<sup>6</sup> I should at this point perhaps mention a distinction that is familiar from the literature, and that one *might* think is similar to that which I have proposed between I- and R-beliefs: namely, the belief-acceptance distinction; for which see, e.g., Bratman 1992, Cohen 1992, Stalnaker 1984 and van Fraassen 1980. Now, there does not seem to be any very widely accepted understanding of this distinction, and so one possibility

I will return to the proposed picture of belief later in the paper. First, however, I want to consider whether one might not be able to solve the puzzle in a less radical way: i.e., without introducing a new picture of belief.

### 3. Solving the Puzzle without the New Picture?

Specifically, I want to consider whether one might not be able to solve the puzzle simply by appealing to *degrees* of belief: i.e., something which people typically think we have anyway; so if we *could* solve the puzzle this way, that would seem to be more economical. Thus, one might think that degrees of belief would allow a relatively straightforward solution, as follows. The basic idea would be to understand binary belief in terms of degrees in something like the following way: in any given situation, there would be some number  $x$  (with  $0 \leq x \leq 1$ ) such that  $S$  counts as believing  $p$  (in that situation) iff she believes  $p$  to degree at least as great as  $x$ . Thus the idea would be that something about  $S$ 's situation determines a certain threshold  $x$ , and  $S$  counts as believing something just in case her degree of belief in that thing is at least as great as  $x$ .<sup>7</sup> Further, the idea would be that what we are willing to use in reasoning is also tied to this threshold  $x$ : so, in particular, we are willing to use in reasoning anything that we believe to degree at least as great as  $x$ . But—the idea would continue—different situations determine different thresholds, and so  $S$  can go from believing to not believing (for example), without changing her underlying degrees of belief. The hope would then be that this would be enough to solve the puzzle. More precisely, since binary belief is now being understood in terms of degrees, one would replace (I) with:

(I') Degrees of belief are our means of storing information.

---

might seem to be to think of this paper as giving a new account of, and a new argument for, *it*. This is probably *not* a helpful way of thinking about things, however: for the proposed distinction between I- and R-beliefs would in fact seem to be quite different from those that have been proposed under this guise. For example, proponents of the belief-acceptance distinction have not given any indication that they would exclude from the belief-category anything like Rachel's attitude to her plans for next summer. Thus—for this reason among others—it would seem best not to try to construe the proposed distinction as a version of the belief-acceptance one.

<sup>7</sup> For accounts of belief along these lines see, e.g., Foley 2009, Hawthorne 2009 and Weintraub 2001; see also Christensen 2004 for a sympathetic presentation of such an account.



And the hope would then be that the threshold account of belief would allow us to hold onto (R) unchanged.

To see how this is supposed to work, consider again the main example we have been discussing. The idea would be that in the first part of the example, when Rachel comes to believe that she will be in the US next summer, she in fact comes to believe this to some degree  $y$  (which is sufficient in this situation for her to count as binary believing it). But (the idea would continue) in the second part of the example what happens is simply that the threshold for belief rises: and so, even though Rachel's degree of belief in the proposition in question remains unchanged, she no longer counts as believing it; and so (R) no longer commits us to saying that she will be willing to use this proposition in her reasoning. But, crucially, we get this change in what Rachel is willing to use in her reasoning *without* saying that her information changes, and so we do *not* get the puzzling result that the information she acquired in the first part of the example is no longer available for her to use in the third part. Thus perhaps degrees of belief (together with this account of binary belief in terms of them) are all that we need to solve the puzzle after all?

Unfortunately, however, it is easy to vary the example so as to reintroduce the problem (i.e., to give a puzzle that this proposed solution cannot be extended to). Thus, suppose that a few years ago a friend of Rachel's went on safari. And suppose that she was driven around by a wonderful guide called Miriam. Suppose, further, that Rachel does not *just* want to go on safari: she wants to go with Miriam. And suppose, finally, that when Rachel is in the deli, rather than thinking, 'if I win, then I will be on safari next summer', she instead thinks: if I win, then I will be able to go on safari with any guide I like this summer; Miriam is a safari guide; so if I win, then I will be able to go on safari with Miriam this summer. (But suppose that everything else in the example is unchanged.)

Now, we are presumably going to want to say that, when Rachel is in the deli, she believes that Miriam is a safari guide: after all, she is willing to use this proposition in her reasoning in just the way that she is willing to use her beliefs, and she would presumably also be willing to assert it (along with the other premises of her reasoning) if she was asked why she decided to buy the ticket. Thus, if  $x$  is the degree to which Rachel believes that Miriam is a safari guide (in this part of the example), then the threshold for belief

here is no greater than  $x$ . The problem for the proposed solution, however, is that we can also plausibly assume that, if  $y$  is the degree to which Rachel comes to believe that she will be in the US next summer, in the first part of the example, then  $x \leq y$ : for, she is really pretty sure that she will be in the US next summer, whereas she is well aware that Miriam could have changed careers, or been taken ill, or moved to New York, etc. But then the puzzle is back—and in a way that the proposed solution is powerless to solve. For Rachel is *not*, in this part of the example, willing to use in her reasoning that she will be in the US next summer. Thus, by (R), together with the fact that the threshold for belief in this situation is  $\leq x \leq y$ , it follows that she no longer believes this proposition to degree  $y$ . That is, we are once again forced to say that Rachel has lost the information she acquired in the first part of the example (since, by (I'), our store of information is now identified with our degrees of belief; and Rachel no longer has the degree of belief she acquired). Thus, it seems that degrees of belief are not enough to solve the problem, after all.<sup>8,9</sup>

#### 4. Back to the New Picture

In this section I will return to the new picture, and try to go some way towards filling in and refining the outline that I gave above.

---

<sup>8</sup> Note that the account proposed in §2 does not have any problem with this variant example: because, in the deli (in the variant example), Rachel *is* taking seriously worlds where she wins the lottery; whereas she is *not* taking seriously worlds where Miriam is no longer a safari guide. As a result, the variant example is handled just as the original was.

<sup>9</sup> An alternative account of binary belief in terms of degrees has been proposed by Brian Weatherson (see 2005 and 2012; similar proposals are made in Ganson 2008 and Fantl and McGrath 2010). The basic idea is that  $S$  believes that  $p$  iff her degree of belief in  $p$  is sufficiently high that changing it to 1 would not change her answer to any relevant question. This account is specifically developed to handle (among other things) beliefs about lotteries, and it should be able to handle the variant example just described. Thus, a fuller presentation of the proposal of this paper would contain a detailed discussion of Weatherson's account. One problem with this account, however, is that (in its simple form) it gives the result that people (almost always) have inconsistent beliefs: for if neither  $p$  nor  $\text{not-}p$  bear on any relevant question, then (on this account) I will count as believing both. Now, Weatherson is well aware of this problem, and he has a patch to propose (see 2005 and 2012; and note that the proponents of similar proposals would *also* seem to require such patches). However, this patch makes the account significantly more complicated, and, thus, less attractive. Further, I would argue that (even independently of this issue) the picture that I have proposed gives a much more straightforward treatment of the examples that generate the puzzle. For a thorough critique of Weatherson's (and similar) proposals, see Ross and Schroeder, forthcoming.

#### 4.1. Information Sets vs Relevance Sets

The first thing that I want to discuss is the division of labour between information sets and relevance sets. For, on the picture that I have proposed, what one is willing to use in reasoning is determined by the intersection of these (by  $(R^*)$ , one is willing to use  $p$  in reasoning if one  $R$ -believes it, and one  $R$ -believes it iff it is true at every world in the intersection of these two sets). But the upshot is that even if we know exactly which propositions  $S$  is willing to use in her reasoning (at a particular moment), this will not necessarily allow us to determine what her information set is and what her relevance set is (because, of course, distinct pairs of sets can have the same intersection); and if one cannot determine what her information and relevance sets are, then neither can one determine which of the propositions she is willing to use in reasoning she  $I$ -believes, and which she merely  $R$ -believes.

Now, the question of how exactly these sets divide the labour is presumably an empirical one. One simple hypothesis might be that one only ever rules out (i.e., excludes from one's information set) worlds at which one's course of experience is different from its actual course. Thus, every world at which my experiences are as they actually are would be in my information set: including worlds where I am a brain in a vat, or where the universe is about to cease to exist, etc. On this hypothesis, then, all of the work of getting from this basis to the things that we actually use in reasoning (e.g., that there is a computer in front of me, that it is raining outside, etc.) would be the job of our various relevance sets.

So that is one hypothesis that one might make about the division of labour. But alternatives, with less austere accounts of our information, are also possible. Thus, one might think that we *also* rule out worlds in the following ways, for example: (i) sense perception (e.g., if I see a computer in front of me, then I would exclude from my information set worlds at which there is *not* a computer in front of me); (ii) a combination of others' sense perception and testimony (e.g., if you see an elephant outside, and tell me about it, then I will exclude from my information set worlds at which there is *not* an elephant outside); and (iii) basic inductive inference (e.g., I might rule out from my information set worlds at which you drop that vase but at which it doesn't break).

Thus, a range of hypotheses would seem to be available. Those with less austere accounts of our information seem to me to be the most plausible. But I will content myself here with these descriptions of some of the alternatives; arguing for one in particular will have to wait for future work.

#### *4.2. More than One Relevance Set at a Time?*

In talking about the new picture, I have implicitly assumed that subjects will at any given moment have no more than one relevance set. This is surely an oversimplification, however. For example, suppose that Rachel is walking to the deli to buy a lottery ticket while at the same time talking on her cell phone about how she is going to spend her next summer in the US.<sup>10</sup> Then, relative to one of her tasks (i.e., walking to buy the ticket), she is *not* willing to use in her reasoning that she will be in the US next summer (for, if she was, then she would conclude that the ticket would be a waste of money, and thus that she should turn around and walk in the opposite direction). But, relative to the other (i.e., her phone conversation), she *is* willing to use this proposition in her reasoning. It would seem that the proposed picture can straightforwardly be extended to cover such multi-tasking, however. One must simply say that relevance sets (and, in turn, R-beliefs) are relative not simply to subjects and times, but, rather, to subjects, times and activities (so if a subject is engaged in *n* activities, she will have *n* relevance sets). So extended, the picture would seem to apply happily to multi-tasking Rachel just as it applied to her earlier incarnation.

#### *4.3. Less than One Relevance Set at a Time?*

But there is also the opposite issue: the idea behind relevance sets is that in different situations I will take different worlds seriously (depending on the sort of cognitive activity I am engaged in); but what about times at which I am not engaged in *any* conscious cognitive activity, and where I would thus not seem to be taking *any* worlds seriously? For example, when I am in a dreamless sleep? Now, it is presumably true that

---

<sup>10</sup> This case is inspired by one presented in DeRose 2009, 269–72.

at such times I will *not* have a relevance set; and thus that I will *not*, strictly speaking, have R-beliefs.<sup>11</sup> Nevertheless, we could surely make perfectly good sense of a practice of ascribing R-beliefs to people at such times: such talk would simply have to be understood as making implicit reference to a certain sort of cognitive activity, and thus to a certain sort of relevance set.

#### *4.4. There is Nothing We are Always Willing to Use in Reasoning!*

A worry one might have about the proposed picture stems from the fact that there does not seem to be *anything* that we are *always* willing to use in reasoning. For example, suppose that I am offered a bet that pays a penny if *p* is true, and that offers an eternity of torture otherwise. Then it would seem that, whatever *p* is, I will *not* in such a situation be willing to use it in my reasoning (to conclude, for example, that I would be better off taking the bet than passing on it). But how is this to be handled on the proposed picture, without saying that in such a case I must have lost the information that *p*? For, on the proposed picture, we are *always* willing to use in our reasoning anything that is true throughout our information set (by (R\*)) together with the definition of R-belief). Thus, there would seem to be little alternative to saying that, if *p* is something that is so true, then: if I am offered such a bet, I must lose the information that *p*. But this conclusion is surely just as unacceptable here as it was in the original puzzle case involving Rachel. Hence the concern!

What I want to suggest, however, is that there is a very natural generalization of the proposed picture that can handle even such cases.<sup>12</sup> For, the basic idea behind this picture is that there are two distinct components to belief: on the one hand, there is our information; and, on the other, there is what we are willing to use in reasoning. In particular, I have thought of our store of information as a set of worlds, and then I have said that, to get at what we are willing to use in reasoning, you have to *remove* from this set any worlds that we are not taking seriously. Perhaps, however, the correct moral of the bet example is that this is an oversimplification: it is not simply that in certain

---

<sup>11</sup> Thus, the principle that I-belief entails R-belief (see footnote 5) strictly speaking only holds at times at which I actually *have* R-beliefs (i.e., at which I actually *have* a relevance set).

<sup>12</sup> Thanks to John Hawthorne here.

situations we *remove* worlds from our information set; rather, we also sometimes *add* them. Thus, in the bet situation, the idea would be that we temporarily add to our information set some worlds at which  $p$  is false. More generally, one would generalize the notion of a relevance *set* to that of a relevance *pair* of (disjoint) sets,  $\langle R^-, R^+ \rangle$ : thus, if  $I$  is our information set, then—the idea would be—we are willing to use in our reasoning anything that is true throughout  $(I \cap R^-) \cup R^+$  (the case in which one *only* has a relevance set  $R$  would then simply be a case in which  $R^- = R$  and  $R^+ = \emptyset$ ). But the basic idea behind the picture would be preserved: for we would still have two components—a stable store of information, together with something more variable—and what we are willing to use in reasoning would still be determined by modifying the former using the latter. Thus, I hope that even the sort of problematic case considered can be handled by a natural generalization of the proposed picture.<sup>13</sup>

I hope in this section, then, to have gone at least some way towards filling in, and refining, the outline that I gave in §2.

## 5. Desire

In this final section, however, I want to consider a different way in which one might try to argue for the proposed picture of belief. In particular, I want to do this via considerations from desire. More precisely, the idea is as follows. I will argue that there are independent reasons for giving an account of desire that is, in important respects, very similar to the proposed picture of belief. Thus, since one would presumably like to give a unified account of propositional attitudes in general—and so of belief and desire in particular—this will amount to a new argument for the picture of belief. The thought behind this last point is simply this: the whole idea behind the general notion of a propositional attitude is that there is something important that a whole range of mental

---

<sup>13</sup> A further reason for wanting to generalize the proposed picture along the lines suggested would be so as to account for things such as reasoning under a hypothesis (for, on the proposed generalization, this could be thought of in terms of temporary additions to our information set). However, for reasons of space I will not try to develop this idea here.

states have in common; and it would thus seem to be theoretically virtuous to take this unity as far as it is possible to.

I want to start, then, by asking some basic questions about desire. The first of these is simply: what exactly does it *mean* to desire that p (for a proposition p)? For example, suppose that I want a ham sandwich (that is, suppose that I desire that I have a ham sandwich): what exactly does this amount to? And the first *point* that I want to make is just this: in general, this seems to be a very intricate matter. For there are all sorts of ways in which this proposition (i.e., the proposition that I have a ham sandwich) could be made true: for example, I could have a delicious, freshly prepared sandwich from the best restaurant in town; or I could have one that is completely tasteless, or rotten, or poisoned, or... Now, to count as desiring this proposition I presumably have to like the idea of at least *some* of these ways it could be made true—but exactly *which* is, *prima facie*, unclear.

However, if one *instead* considers desire-like attitudes to whole worlds (for example, preferences between these) then there would seem to be no comparable intricacy. For there are *not* multiple ways for a world to be actual, in the way in which there *are* multiple ways for a proposition to be true. As a result, while it is hard to say what exactly desiring that p amounts to (or preferring p to q, for that matter), there is no comparable unclarity when it comes to the question of what preferring one *world* to another amounts to.

But this then suggests that a very natural strategy for giving an account of desire would seem to be this: take desire-like attitudes to worlds as basic, and give an account of desiring that p in terms of these. But this is then already to say that we should give an account of desire that is in a fundamental respect similar to the proposed picture of belief: for, in each case, we would be taking attitudes to worlds as basic, and giving an account of the propositional attitude in question in terms of these (in the belief case the attitudes to worlds we took as basic were of course ruling out and taking seriously, and we gave an account of (two sorts of) propositional belief in terms of these). What I now want to argue, however, is that once one thinks in a bit more depth about what one's account of desire should look like, then even greater similarities with the proposed picture of belief will emerge.

Thus, I now want to try to think in a bit more detail about how exactly one's account of desire might go. So suppose that one takes preferences between whole worlds as basic: how might one give an account of desiring that *p* in terms of these?

Since desiring that *p* presumably in *some* sense involves preferring *p*-worlds to not-*p*-worlds (i.e., preferring worlds at which *p* is true to worlds at which it is not), a very natural first thought is that desiring that *p* amounts simply to preferring *any* *p*-world to *any* not-*p*-world. Unfortunately, though, this cannot quite be right (and for, essentially, reasons we have already considered). For I can surely desire a ham sandwich without its being the case that I prefer *all* worlds at which I have a poisoned ham sandwich to *all* worlds at which I do not have a ham sandwich at all. But that means that the proposed account of desire is not right.

But a very natural *second* thought is then that desiring that *p* amounts to one's 'most preferred' worlds all being *p*-worlds: thus—the idea would be—it is not required that one prefer *every* *p*-world to *every* not-*p*-world; but there must be some group of 'most preferred' worlds, all of which must be *p*-worlds. Thus, the account of desire would look something like this: *S* desires that *p* iff there is some world *w* such that every world that *S* likes at least as much as *w* is a *p*-world; so, intuitively, the idea would be that there is some point in *S*'s preference ranking (represented by *w*) such that every world at or above that point is a *p*-world.

Unfortunately, though, this *still* isn't quite right. For suppose that I am in a deli choosing which sort of sandwich to order. Surely I can truthfully say 'I want a ham sandwich'—and surely I can count as desiring that I have a ham sandwich—even if I would much prefer to have caviar; i.e., even if there are caviar-worlds that I prefer to *all* ham-sandwich-worlds.

A natural diagnosis of what has gone wrong is this: the reason that you count as desiring a ham sandwich, in such a situation, despite your preference for caviar-worlds, is that caviar-worlds are *irrelevant* here (i.e., you are not taking them seriously). A natural *third* thought is thus this: *S* desires that *p* iff there is some world *w* such that every *relevant* world that *S* likes at least as much as *w* is a *p*-world. This would then seem to get around the 'caviar problem': for, in the deli, I am not taking caviar-worlds seriously, and so they do not give a counterexample to the proposed account. And I would suggest



that, more generally, what we have arrived at would seem to be a very plausible account of desire.<sup>14,15</sup>

But, if this account of desire is on the right track, then we are going to end up with something that looks a lot like the proposed picture of belief. For in each case there is both a stable component and a more variable relevance set (thus, in the belief case the stable component is our information set, while in the desire case it is our preference ranking of worlds). Further, in each case this pair of components will give rise to a pair of species of the attitude in question—a stable and a variable one. Thus, in the belief case, the two species arise as follows: one stably (i.e., I-)believes something iff it is true at every member of one's information set; whereas one variably (i.e., R-)believes it iff it is true merely at every *relevant* world in one's information set. Similarly, in the desire case, one will get two species of the attitude as follows.<sup>16</sup> On the one hand, there will be 'stable desires': where S 'stably desires' that p iff her most preferred worlds are all p-worlds (i.e., iff there is some w such that p is true at every world that she likes at least as much as w). But, on the other hand, there will also be 'variable desires': where S 'variably desires' that p iff her most preferred *relevant* worlds are all p-worlds (i.e., iff there is some world w such that p is true at all of the *relevant* worlds that she likes at least as much as w). Thus, to return to our example: while I might variably desire a ham sandwich, I will

---

<sup>14</sup> This account of desire is similar in structure to accounts of various modal terms that have been proposed by Angelika Kratzer (see, e.g., 1991). Insofar as Kratzer's proposals have been successful, that perhaps offers *some* support for the proposed account of desire.

<sup>15</sup> An alternative account of desire in terms of possible worlds has been proposed by Irene Heim (see 1991). The proposal is essentially this: S desires (or wants) that p iff for every doxastically possible world w, (i) if w is a p-world then S prefers w to its nearest not-p-alternatives; and (ii) if w is a not-p-world then S prefers w's nearest not-p-alternatives to w. Unfortunately, however, this account seems prone to counterexamples. For example, suppose that I am at a wedding and I have to choose ice cream or cheese. If I choose ice cream, then I will randomly be assigned one of ten varieties, and similarly if I choose cheese. Suppose, further, that I prefer nine of the flavours of ice cream to all of the types of cheese, but that I prefer all of the types of cheese to the other flavour (coffee, say, which I detest). (And suppose that I know all of this about the set-up.) It is then very natural for me to say, 'I want ice cream, but I don't want coffee flavour (and so I'm not sure what to do)'. But, on Heim's proposal, I do *not* count as wanting ice cream, because there are doxastically possible ice-cream-worlds that I do *not* prefer to their nearest not-ice-cream-alternatives (because I do not prefer coffee-flavour-worlds to their nearest not-ice-cream-alternatives, i.e., worlds at which I have one of the varieties of cheese).

<sup>16</sup> In my initial presentation in this section, I have only made the case for thinking of desire as the variable attitude here; whereas, in the belief case, the puzzle made cases both for the identification of belief with I-belief, and *also* for the conflicting identification with R-belief. A fuller discussion of desire, however, would argue that there is a similar tension in this case (using examples similar to those that generated the puzzle for belief).

typically *not* stably desire one (because there is always *something* that I would prefer; if not caviar then...).

Thus, there would seem to be independent reasons for giving an account of desire that is, in important respects, similar to that which I proposed for belief. And this would then seem to constitute a further argument for the account of belief itself: for one would surely like to give a unified account of these attitudes, as opposed to saying that we have two very different cognitive capacities, one of which is responsible for belief, and the other for desire.

Thus, I hope to have made a strong case for a new picture of belief: a picture on which information and reasoning each get their own kind.<sup>17,18</sup>

## References

- Bratman, M. 1992. 'Practical Reasoning and Acceptance in a Context'. *Mind* 101: 1–15.
- Christensen, D. 2004. *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford: Oxford University Press.
- Cohen, J. L. 1992. *An Essay on Belief and Acceptance*. Oxford: Oxford University Press.
- Cohen, S. 1988. 'How to be a Fallibilist'. *Philosophical Perspectives* 2: 91–123.
- DeRose, K. 2009. *The Case for Contextualism: Knowledge, Skepticism, and Context, Vol. 1*. Oxford: Oxford University Press.

---

<sup>17</sup> I would like to close by just mentioning an alternative way in which this paper might have been written; i.e., an alternative way in which the basic ideas might have been presented. What this would have emphasized is what these ideas tell us about the way in which our minds store information. Now, the *prima facie* natural view about this would seem to be that we store information in the form of propositions (where these are understood in *something like* the Fregean or Russellian mould): for, it would seem that we store information by having beliefs; and thus that our store of information amounts to the set of propositions we believe. But what this version of the paper would argue is that we in fact store information in the form of worlds. For, the main example of §1 above would seem to show that the propositional account cannot be right: for, given the plausible principle that we are willing to use in reasoning what we believe, it would give the (unacceptable!) result that Rachel loses the information about what she is going to do next summer as soon as she starts thinking about lottery tickets. The solution, I would then argue, is a new picture of mental information on which our store of information is not a set of propositions, but a set of worlds (i.e., the picture of §2 above). Of course, the idea that we store information in the form of worlds has (in effect) been argued for before (e.g., by Stalnaker 1984). However, this alternative version of the paper would seem to give a new, and very simple, argument for this position. But I leave it to the reader to decide which version would have been better!

<sup>18</sup> For comments and discussion I am extremely grateful to David Chalmers, Keith DeRose, Adam Elga, Katrina Elliott, Branden Fitelson, Elizabeth Harman, John Hawthorne, Josh Knobe, John Morrison, Ted Sider, Zoltán Gendler Szabó, Tim Williamson, and the members of a seminar at Yale.

- Fantl, J. and M. McGrath. 2010. *Knowledge in an Uncertain World*. Oxford: Oxford University Press.
- Foley, R. 2009. 'Beliefs, Degrees of Belief, and the Lockean Thesis'. In *Degrees of Belief*, edited by F. Huber and C. Schmidt-Petri, 37–47. Berlin: Springer.
- Ganson, D. 2008. 'Evidentialism and Pragmatic Constraints on Outright Belief'. *Philosophical Studies* 139: 441–58.
- Harman, G. 1973. *Thought*. Princeton: Princeton University Press.
- Hawthorne, J. 2009. 'The Lockean Thesis and the Logic of Belief'. In *Degrees of Belief*, edited by F. Huber and C. Schmidt-Petri, 49–74. Berlin: Springer.
- Hawthorne, J. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Heim, I. 1992. 'Presupposition Projection and the Semantics of Attitude Verbs'. *Journal of Semantics* 9: 183–221.
- Kratzer, A. 1991. Modality. In *Semantics: An International Handbook of Contemporary Research*, edited by A. von Stechow and D. Wunderlich, 639–50. Berlin: de Gruyter.
- Lewis, D. 1996. 'Elusive Knowledge'. *Australasian Journal of Philosophy* 74: 549–67.
- Ross, J. and M. Schroeder. Forthcoming. Belief, Credence, and Pragmatic Encroachment. *Philosophy and Phenomenological Research*.
- Stalnaker, R. C. 1984. *Inquiry*. Cambridge, MA: MIT Press.
- van Fraassen, B. C. 1980. *The Scientific Image*. Oxford: Oxford University Press.
- Vogel, J. 1990. Are there Counterexamples to the Closure Principle? In *Doubting: Contemporary Perspectives on Skepticism*, edited by M. Roth and G. Ross, 13–27. Dordrecht: Kluwer.
- Weatherson, B. 2005. 'Can We Do without Pragmatic Encroachment?' *Philosophical Perspectives* 19: 417–43.
- 2012. Knowledge, Bets, and Interests. *Knowledge Ascriptions*, edited by J. Brown and M. Gerken, 75–103. Oxford: Oxford University Press.
- Weintraub, R. 2001. 'The Lottery: A Paradox Regained and Resolved'. *Synthese* 129: 439–49.