

## II

## On the Aim of Belief

## Introduction

There are several reasons for being interested in the fact—if it is a fact—that belief aims at the truth. I am going to argue that it's a fact. But first, the reasons for taking an interest in it.

*Reasons for studying truth-directedness*

One reason, first pointed out by Bernard Williams, is that belief's aiming at the truth enables us to explain the difficulty of believing at will.<sup>1</sup> The explanation is that believing a proposition at will would entail believing it without regard to whether it was true, and hence without the aim requisite to its being a belief.

Another reason for taking an interest in the truth-directedness of belief is that it may explain the phenomenon that has generally come to be known as the normativity of content. It may even explain the phenomenon away.

The phenomenon to be explained in this case is that the attribution of propositional content to a thought seems to yield normative consequences, on

In writing this chapter I have benefited from conversations with Linda Brakel and Nishiten Shah. I have also received helpful comments and suggestions from Paul Boghossian, Michael Bratman, John Broome, Jennifer Church, Stephen Everson, Tamar Gendler, John Gibbons, David Hills, Paul Horwich, Jim Joyce, Mike Martin, Ruth Garrett Millikan, Richard Moran, Jerome Neu, Lucy O'Brien, David Papineau, Georges Rey, Gideon Rosen, Zoltan Szabo, Stephen Schiffer, Peter Vranas, and Ken Walton. The chapter was presented to the 1998 Chapel Hill Colloquium, with comments by Gideon Rosen; to a conference at University College London, with comments by Lucy O'Brien; to a Philosophy of Mind seminar at New York University conducted by Paul Boghossian and Stephen Schiffer; and to the Philosophy Departments of Stanford University, Washington University, the University of Missouri, Columbia, and Cornell University. Work on this chapter was supported by a fellowship from the John Simon Guggenheim Memorial Foundation, together with matching grants from the Department of Philosophy and the College of Literature, Science, and the Arts, University of Michigan.

<sup>1</sup> 'Deciding to Believe,' in *Problems of the Self* (Cambridge: Cambridge Univ. Press, 1973), 136–51. I have argued elsewhere that belief's aiming at the truth does not prevent us from adopting a belief at will if our adopting it can be expected to make it true (*Practical Reflection* (Princeton: Princeton Univ. Press, 1989), 127 ff.); see also Chap. 8, above, pp. 194–6.

purely conceptual grounds. To say of a thought that it has the content that snow is white, for example, seems to imply that one ought to have the thought only if snow is white; and this normative consequence seems to follow immediately from the very concept of a thought's possessing such a content. The normative force inherent in content attributions has struck some philosophers as evidence against their being analyzable in naturalistic terms or assimilable into the natural sciences.<sup>2</sup>

But the normative implications claimed here for content attributions are not quite right. It isn't true that one ought to have the thought that snow is white only if snow is white; or, at least, it isn't true on the most inclusive interpretation of the phrase "to have a thought." If one can have a thought merely by entertaining it, without belief, then having the thought that snow is white would be perfectly in order even if one were up to one's neck in black snow. What one would be obliged to avoid, if snow weren't white, is not the mere thought of snow's being white but rather the belief with that content. And one would be obliged to avoid the belief because, given its nature as a belief, it would aim at being true but, given its content, would fall short of this aim on account of snow's not being white. The belief would be wrong or incorrect in the sense that it would be a failure in relation to its own aim.

The normativity previously attributed to content thus turns out to arise from the combination of content plus belief. We may even be tempted to say that the normativity is due entirely to the truth-directedness of belief and not at all to the nature of content. All that follows from a thought's having the content "snow is white" is that the thought is true if and only if snow is white. Whether the thought ought or ought not to be held in a world containing black snow depends on whether it is to be held in a way that aims at the truth—that is, as an object of belief.

The result appears to be that in order to naturalize the normativity associated with content, we need only naturalize the truth-directedness of belief. If it can be a natural or scientific fact that belief aims to be true, then it can also be a natural or scientific fact that false beliefs are wrong or incorrect, which is the fact underlying the normativity generally attributed to content.<sup>3</sup> The hope of naturalizing that normativity is thus a reason for being interested in how belief aims at the truth.

A third reason for taking an interest in the truth-directedness of belief is that it may help us to understand theoretical reasoning and perhaps, by analogy,

<sup>2</sup> See Saul A. Kripke, *Wittgenstein on Rules and Private Language* (Cambridge, Mass.: Harvard Univ. Press, 1982).

<sup>3</sup> Naturalizing the aim of belief is not on my agenda for this paper. For an attempt to naturalize the normativity of content via the nature of belief, see Ruth Millikan, 'Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradox,' in *White Queen Psychology and other Essays for*

X  
belief  
No. Norm  
see account  
in nature of  
attitude

?

✓

practical reasoning as well. Theoretical reasoning justifies a belief by adducing considerations that indicate it to be true. And a belief can be justified by indications of its truth because being true is what would make it successful or correct, given that being true is its aim as a belief. Hence the truth-directedness of belief is what accounts for the justificatory force of theoretical reasoning.

I have argued on other occasions that practical reasoning must be understood along similar lines.<sup>4</sup> If there is to be a mode of reasoning that justifies action, as theoretical reasoning justifies belief, then there must be a criterion of correctness or success in relation to which action can be justified, just as there is a criterion of correctness or success for belief. If there's nothing that constitutes success for action as such, in the way that truth constitutes success for belief, then practical conclusions will not be supported by considerations with the normative force of justification or reasoning.

I am not going to defend this argument here—nor either of the preceding arguments, for that matter. My goal in this chapter is not to prove that the truth-directedness of belief can help us to understand the difficulty of believing at will, or the normativity generally attributed to content, or the form of practical reasoning. I have introduced these topics only as ulterior reasons for being interested in the phenomenon that I want to discuss.

### Summary

I'll begin the next section by teasing apart the multiple relations that belief bears to the truth, setting aside those relations which do not amount to belief's having the truth as its aim. I will argue that the sense in which belief aims at the truth is not that it involves believing a proposition to be true; nor that it involves regarding a proposition *as* true; but rather that it involves so regarding a proposition with a particular aim. Just as a proposition is assumed by being regarded as true for the sake of argument or testing, so a proposition is believed by being regarded as true for the sake of something else. That "something else"—the aim with which a proposition must be regarded as true in order to be believed—is the aim of getting its truth-value right, by regarding it as true only if it really is. However, I will briefly postpone defining what it is to regard a proposition as true. Instead, I'll close the section by identifying various ways in which this subdoxastic attitude of acceptance, as I'll call it, can satisfy the concept of having an aim, or how its aim can be realized in the subject's psychology.

At the start of the third section, I'll define regarding-as-true as the attitude that plays a particular motivational role—indeed, the role that is traditionally

*Alice* (Cambridge, Mass.: MIT Press, 1993), 211–39. I would favor a strategy that is only slightly different in emphasis from Millikan's: see n. 18, below. <sup>4</sup> See Chaps. 5, 8, and 10, above.

thought to be definitive of belief. I will thus commit myself to the claim that this motivational role, far from being definitive of belief, is definitive instead of the subdoxastic attitude of acceptance, which is involved in assuming, as I've just mentioned, as well as other cognitive attitudes, such as imagining. I'll devote most of this section to arguing that the attitude of imagining does indeed play this motivational role. My conclusion will be that belief must be characterized, not just as the attitude having the motivational role, but rather as a truth-directed species of that attitude: to believe a proposition is to regard it as true with the aim of thereby accepting a truth. The final section will deal with a few objections to this conclusion.

### What is Truth-directedness?

The statement that belief aims at the truth is supposed to reveal something about the nature belief, something that distinguishes it from other propositional attitudes. So if we want to define truth-directedness, perhaps we should start by asking what is distinctive about the nature of belief. Not all relations between belief and the truth are sufficiently revealing.

### What truth-directedness isn't: believing-true

For example, every instance of believing is an instance of believing something to be true, and this relation to the truth is sometimes confused with truth-directedness.<sup>5</sup> But in bearing this particular relation to the truth, belief is just like any other propositional attitude, since wishing entails wishing something to be true, hoping entails hoping something to be true, desiring entails desiring something to be true, and so on. Hence the fact that believing entails believing-true doesn't set belief apart from other attitudes, as truth-directedness is supposed to do.<sup>6</sup>

<sup>5</sup> For this confusion see, e.g., David Gauthier, 'Assure and Threaten,' *Ethics* 104 (1994) 699. Zoltan Szabo has persuaded me that there may be a species of believing that doesn't entail believing-true. If there were, my discussion in this section would have to be modified, but its overall point wouldn't be undermined.

<sup>6</sup> Steve Schiffer has pointed out that, although every propositional attitude  $\phi$  is such that it entails  $\phi$ ing-true, belief is the only attitude  $\phi$  that entails believing-true. (Or, more precisely, it's the only non-composite attitude that entails believing true. Some composite attitudes also entail believing-true—for example, regretting.) Although the property of being a non-composite attitude that entails believing-true does distinguish belief from other attitudes, it does so without conveying much about the nature of belief. My working assumption in this paper is that truth-directedness is not just distinctive of belief but informatively so. (See also n. 8, below.) What would be most informative, of course, is a fully reductive account of belief, in terms of uncontroversially naturalistic concepts. I do not claim to have a fully reductive account (but see n. 16, below).

One might take exception to the last item on the foregoing list: maybe we can simply desire an object, without desiring anything to be true. Similarly, one might claim that we can simply imagine an object, without imagining anything to be true.<sup>7</sup> One might then draw a contrast with belief by pointing out that, although there are objects that we believe without believing *them* to be true (as when we believe a person, for example), we still cannot believe them unless there is *something* that we believe to be true (such as something the person has told us).

Yet the contrast that one would have drawn, in that case, is a contrast between belief and attitudes that are thought to have non-propositional instances. That is, one would have claimed that there are instances of desiring and imagining that don't entail desiring or imagining that *p*, for some proposition *p*, whereas there cannot be non-propositional instances of believing. In saying that believing entails believing-true, one would have said no more than that belief must have a propositional object, and so one still wouldn't have said anything that distinguished believing from propositional instances of desiring and imagining. For in cases of desiring or imagining that *p*, the proposition *p* is desired or imagined to be true, just as a proposition is believed to be true whenever it is believed.

The point is that truth-directedness is supposed to distinguish belief from other attitudes generally, and not just from their non-propositional instances. That believing entails believing-true cannot be what is meant, then, by the observation that belief aims at the truth.

#### *What truth-directedness isn't: regarding-as-true*

We can make some progress toward distinguishing belief from other attitudes by noting that the similarity between the expressions 'believing to be true' and 'desiring to be true' conceals an underlying difference. Believing a proposition to be true entails regarding it as something that is true, as a truth already *in*

<sup>7</sup> Sebastian Gardner claims that the unconscious phantasies cited in psychoanalytic explanations are non-propositional (*Irrationality and the Philosophy of Psychoanalysis* (Cambridge: Cambridge Univ. Press, 1993), 104, 122, 155–6, 189–91). The object of such phantasies wouldn't be a truth-bearer, and so they wouldn't entail imagining-true. Although I believe that some imagining doesn't entail imagining-true, I don't believe that the reason is necessarily that it is non-propositional. Such imagining may simply consist in picturing a state of affairs without engaging in the mental fiction *that* it obtains. (See n. 11, below.) I agree that some imagining may not even be propositional—for example, imagining a particular color without imagining any state of affairs involving it. But such rudimentary cases of imagining wouldn't do the work required of the phantasies discussed by Gardner. I suspect that Gardner confuses the question whether unconscious phantasies are propositional with the question whether they are sentential, in the sense of being represented in a language of thought, or readily expressible in ordinary language. They may be non-sentential, in this sense, while also having truth-conditions and hence being propositional.

being; whereas desiring a proposition to be true entails regarding it as something to be made true, as a truth-*to-be*. So in 'believing to be true', the infinitive 'to be true' takes the place of what would be, in direct speech, a predication of truth in the indicative mood; whereas in 'desiring to be true', the infinitive takes the place of what would be a predication of truth in the optative or gerundive.

This distinction is so subtle as to seem like a quirk of grammar rather than evidence of a genuine difference between desire and belief. We must therefore try to flesh out the distinction.

One way *not* to flesh it out is by incorporating indicative and optative predications of truth into the propositional objects of these attitudes. Believing that *p* entails believing *p* to be true, but it does not involve an attitude toward the proposition '*p* is true'; it involves only an attitude toward the proposition *p*. If it did involve an attitude toward the proposition '*p* is true', then the latter attitude would have to be a belief: what other attitude toward '*p* is true' could constitute believing *p* to be true? The result would be a problematic regress.<sup>8</sup> The belief that *p* would involve a belief that *p* is true, which would have to involve a belief that '*p* is true' is true, and so on, *ad infinitum*. Surely, we should say that believing *p* to be true involves, not an attitude toward the proposition '*p* is true', but only an attitude toward *p* itself, albeit an attitude toward it *as* true.

From the fact that believing entails believing-true, we have now derived two features of belief: belief always takes a propositional object, and it regards that object as true. Unfortunately, these two features do not take us very far toward distinguishing belief from the other attitudes.

One respect in which they fall short is that talk of regarding a proposition as true is no more informative the other expressions with which we're already dissatisfied. What is it to regard a proposition as true?

In past discussions, I have refused to answer this question, or I have given a non-answer.<sup>9</sup> My non-answer has been that a complete theory of belief will owe us an account of what it is to regard a proposition as true, but that

<sup>8</sup> What's problematic about the regress would depend on how belief-attributions were interpreted. If the truth of "*S* believes that *p*" required *S* to have a mental representation with the content *p*, then the regress would be vicious, since it would require the subject to have an infinite number of mental representations in order to have even one belief. If "*S* believes that *p*" wasn't interpreted as crediting *S* with a distinct representation, then the regress would be benign, but the resulting account of belief would be uninformative. In that case, "*S* believes that *p*," "*S* believes that *p* is true," "*S* believes that '*p* is true' is true," and so on, would be logically redundant attributions of one and the same attitude; and the fact that the first attribution entails all the rest would just be the trivial fact of their redundancy rather than a substantive fact about the nature of belief. (See also n. 6, above.)

<sup>9</sup> See above, Chap. 8, p. 182, n. 35.

different theories will discharge this obligation differently, offering alternative accounts among which I can reasonably wish to be neutral. I have simply insisted that believing, whatever it turns out to be, will have to involve regarding a proposition as true, whatever *that* turns out to be.

Yet there is another respect in which we have fallen short of distinguishing belief from the other attitudes. Whatever regarding-as-true turns out to be, it will still be involved in more than believing, since it will be involved, for example, in supposing or assuming, and in propositional imagining as well.<sup>10</sup> These attitudes are cognitive, like belief, rather than conative, like desire. To imagine that *p* is to regard *p* as describing how things are, not as prescribing how they should be.<sup>11</sup> Imagining is therefore a way of regarding a proposition as true—or, to introduce a term, a way of accepting a proposition.<sup>12</sup> The question remains how belief differs from imagining and the other cognitive attitudes.

#### What truth-directedness is

But here, at last, is the payoff to our strategy of asking what is distinctive about belief: we have found the role of truth-directedness. Differences among the cognitive attitudes appear to consist in the aim with which they accept a propo-

<sup>10</sup> For discussions that support this claim, see Robert C. Stalnaker, *Inquiry* (Cambridge, Mass.: MIT Press, 1984), 79 ff.; and Michael Bratman, 'Practical Reasoning and Acceptance in a Context,' *Mind* 101 (1992) 1–15, reprinted in Bratman, *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge Univ. Press, 1999), 15–34.

<sup>11</sup> This issue is complicated by an ambiguity in the verb 'to imagine', which can be used to describe a thought either as imagistic in its intrinsic character or as fictional in its intent. In the former sense, imagining may entail no more than entertaining a thought; only in the latter sense does imagining entail regarding the thought as true. The differences between these senses can be marked by a difference in the grammatical form of the complement phrase. To imagine *the moon's being made of green cheese* may simply be to entertain a thought in the form of a mental image, an image of a green-cheesy moon. But to imagine that *the moon is made of green cheese* is to engage in a mental fiction, which may or may not involve imagery. I shall be using the verb 'to imagine' exclusively in the latter sense. (See Kendall Walton's discussion of imagining in *Mimesis as Make-Believe; on the Foundations of the Representational Arts* (Cambridge, Mass.: Harvard Univ. Press, 1990), 19–21.)

<sup>12</sup> The distinction I am drawing here is, in my view, the distinction that philosophers have generally been trying to pinpoint with the term 'direction of fit'. Unfortunately, many discussions of direction of fit tend to conflate regarding-as-true with aiming at the truth. See, e.g., Humberstone, 'Direction of Fit,' *Mind* 101 (1992) 59–83. Because of this confusion, I will avoid the term 'direction of fit' in this paper. I discuss the issue in Chap. 5.

Note also that my use of the term 'acceptance' is stipulative and idiosyncratic. The distinction that I draw between belief and acceptance is very different from distinctions drawn by Keith Lehrer in 'The Gettier Problem and the Analysis of Knowledge,' in *Justification and Knowledge*, ed. George S. Pappas (Dordrecht: D. Reidel, 1979), 65–78; John Perry, 'Belief and Acceptance,' *Midwest Studies in Philosophy* 5 (1980) 533–542; Bas C. Van Fraassen, *The Scientific Image* (Oxford: Clarendon Press, 1980); and L. Jonathan Cohen, *An Essay on Belief and Acceptance* (Oxford: Clarendon Press, 1992). My distinction is slightly different from those drawn by Stalnaker, *Inquiry*, 79 ff.; and Bratman, 'Practical Reasoning and Acceptance in a Context.'

sition, or regard it as true. Assuming, for example, involves accepting a proposition for the sake of argument, or for some similar purpose, but it doesn't involve believing the proposition. When a mathematician says, "Suppose that  $\sqrt{2}$  is rational," he is not inviting us to believe it, but he is inviting us to take an attitude that's more like a belief than it is like a desire. I suggest that this attitude is like a belief because it is an acceptance, and that it is unlike a belief because it is acceptance for the sake of argument, whereas belief is acceptance for the sake of something else.

What else could we accept a proposition for? What purposes or aims could acceptance have? Well, imagining involves regarding a proposition *as* true irrespective of whether it *is* true—regarding it as true, that is, without trying to get its truth-value right. Perhaps, then, believing involves regarding a proposition as true with the aim of so regarding it only if it really is.<sup>13</sup> Thus, to believe a proposition is to accept it with the aim of thereby accepting a truth.<sup>14</sup>

The result is a fairly modest conception of how belief aims at the truth. This conception is modest, to begin with, because it avoids burdening belief with some of the more ambitious aims that might be suggested by talk of truth-directedness.

Perhaps the most ambitious aim suggested by such talk would be the aim of believing as many truths as possible. This aim would be irrational, of course, since the world is teeming with truths, most of which are too trivial to be worth believing. But there are qualified versions of this aim that might not be irrational, such as the aim of believing as many as possible of those truths which are useful or interesting. Also rational, perhaps, would be the aim of maximizing the proportion of truths to falsehoods among one's beliefs.

Yet none of these aims would answer to what I have in mind when postulating an aim that distinguishes belief from other cognitive attitudes. What I have in mind is an aim that differentiates believing a particular proposition

<sup>13</sup> Walton has argued that imagining, though not constrained by the truth, is not entirely unconstrained, either (39 ff.). Various imaginative games and projects may lay down rules for what is appropriate to imagine—for example, a game in which the participants agree to imagine that the tree-stumps in the forest are bears. What is appropriate to imagine in such a context, Walton describes as being "fictional" in that context. And he therefore concludes, "Imagining aims at the fictional as belief aims at the true" (41).

I don't think that this statement is meant to imply that aiming at the fictional is constitutive of imagining. At most, I think, it is meant to imply, conversely, that being aimed at by some form of imagining is constitutive of the fictional. Thus, Walton does not mean to rule out the possibility of aimless imaginings, as I would rule out the possibility of non-truth-directed beliefs. He simply means to define the fictional as that at which some form of imagining aims, in the way that belief aims at the truth. So interpreted, Walton's statement is compatible with my view.

<sup>14</sup> Here and throughout this paper, I deal exclusively with full belief—i.e., belief of the all-or-nothing kind. For the sense in which partial belief aims at the truth, see James M. Joyce, 'A Non-pragmatic Vindication of Probabilism,' *Philosophy of Science* 65 (1998) 573–603.

from other ways of regarding that proposition as true, such as imagining or supposing it. I thus have in mind an aim with which a particular proposition can be accepted, such that its being accepted with that aim constitutes its being believed. And one's acceptance of a proposition can amount to a belief without being part of any global epistemological project of accumulating true beliefs.<sup>15</sup>

What distinguishes believing a proposition from imagining or supposing it is a more narrow and immediate aim—the aim of getting the truth-value of that particular proposition right, by regarding the proposition as true only if it really is. Belief is the attitude of accepting a proposition with the aim of thereby accepting *a* truth, but not necessarily with any designs on truths in general, or Truth in the abstract.

#### *How is the aim realized?*

A further note of modesty in this conception of belief is that it leaves open how the aim of belief is realized. It allows but does not require the aim of belief to be an aim on the part of the believer; and it allows but does not require the aim of belief to admit of a naturalistic reduction. The conception merely requires that belief aim at the truth in some way or other, there being a broad spectrum of ways in which it might do so.<sup>16</sup>

At one end of the spectrum is the case in which a person intentionally aims a belief at the truth, by forming it in an act of judgment. He entertains a question of the form '*p* or not *p*?', wanting to accept whichever disjunct is true; to that end he accepts one or the other proposition, as indicated by evidence or argument; and he continues to accept it only so long as he receives no evidence or argument impugning its truth. The resulting cognition qualifies as a belief because of the intention with which it is formed and subsequently maintained by the believer, and because of the way in which that intention regulates its formation and maintenance.<sup>17</sup>

<sup>15</sup> This issue is discussed by Peter Railton in 'Truth, Reason, and the Regulation of Belief,' *Philosophical Issues* 5 (1994) 73–4.

<sup>16</sup> The following discussion may give the false impression that I hope to reduce the concept of truth-directedness to other concepts that are not teleological. Although I do think that truth-directedness is reducible to some extent, I do not think that its reduction can dispense with teleology. Any reduction will have to allude either to the subject's aims or to the design of his cognitive systems, both of which are teleological notions. I do not regard the ineliminability of teleology to be a flaw in my account of belief; on the contrary, I think of it as a virtue. For I think that the concept of belief just is a teleological concept, of a mental state constituted in part by its point or constitutive aim.

<sup>17</sup> Note that this belief may consist in a cognition that is *misdirected* at the truth. When the subject intentionally sets out to accept whichever proposition is true, he will be guided by methodological beliefs about how to discriminate truth from falsehood. Even if he is wrong about how to arrive at the truth, he is still aiming to arrive at it; his acceptance of a proposition is still regulated in ways that he regards as truth-conducive; and so it will still qualify as a belief, whether or not it is regulated in a manner that is truth-conducive in fact.

A person can also aim cognitions at the truth without necessarily framing intentions about them. Suppose that one part of the person—call it a cognitive system—regulates some of his cognitions in ways designed to ensure that they are true, by forming, revising, and extinguishing them in response to evidence and argument. Regulating these cognitions for truth may be a function for which the system was designed by natural selection, or by education and training, or by a combination of the two.<sup>18</sup> In any case, the system carries out this function more or less automatically, without relying on the subject's intentions for initiative or guidance. Even so, the subject may identify with this system, by endorsing it or fostering it or doing something else that makes its operations attributable to him, in the manner made familiar by Harry Frankfurt.<sup>19</sup> Its workings may then count as his doings, so that he can be said to have regulated the resulting cognitions, and thus to have aimed them at the truth.

At the far end of the spectrum, we can imagine a subject who is dissociated from the workings of this cognitive system, also in Frankfurt's sense: he is oblivious to it, or he disapproves of it, wishes it would stop, hopes it will fail, either in general or on a particular occasion. In that case, the resulting cognitions may not qualify as having been regulated by him. But they will still have been regulated for truth, and hence aimed at the truth, albeit by a part of him with which he doesn't identify. They will still be attempts at accepting truths, even though they will be attempts on the part of a cognitive system rather than the person as a whole. As cognitions aimed at the truth, they will still qualify as beliefs, according to my conception.<sup>20</sup>

<sup>18</sup> This sentence suggests one strategy for naturalizing the truth-directedness of belief. The strategy suggested is slighted different from that pursued by Ruth Millikan in her 'Truth Rules' paper; and it would apply to only some of the mechanisms that regulate beliefs.

The difference in strategy this. On the one hand, beliefs guide the subject's behavior in a manner that benefits him only—or, at least, most reliably—when they are true. Their guiding the subject when true is what confers advantages on him, and so it appears to be what beliefs were selected for, in the course of evolution. Beliefs were thus, metaphorically speaking, designed to be true. On the other hand, beliefs are regulated by psychological mechanisms designed to ensure that they are true. These mechanisms dispose the subject to form and revise his beliefs in response to indicators of their truth—that is, in response to reasons. Thus, beliefs perform their function best when true, just as various bodily systems perform best at 98.6 degrees; and beliefs are regulated so as to be true, just as the body's temperature is regulated so as to be 98.6. I would emphasize the latter explanation as conveying the sense in which belief aims at the truth; Millikan would emphasize the former.

In any case, an evolutionary explanation can apply only to some but not all of the mechanisms that regulate beliefs. Although some of our cognitive mechanisms are designed by evolution, others consist in acquired habits of mind or learned methodologies. The aim of the latter mechanisms may be implicit in the way they are acquired and refined, or explicit in the instruction from which they are learned. As I have said before, naturalizing the aim of belief, or the proper function of belief-regulating mechanisms, is not the purpose of this paper.

<sup>19</sup> *The Importance of What We Care About* (Cambridge: Cambridge Univ. Press, 1988), especially chs. 2, 5, 7, and 12.

<sup>20</sup> I discuss this issue at greater length in Chap. 8, pp. 184–8, above.

All of the cases in this spectrum can be described as follows. An acceptance has the aim of being the acceptance of a truth when it is regulated, either by the subject's intentions or by some other mechanisms in ways designed to ensure that it is true.<sup>21</sup>

Finally, my conception of belief is modest in that it doesn't require belief to be governed by truth-seeking mechanisms alone. There are probably psychological mechanisms that cause, and are designed to cause, beliefs that happen to diverge from the truth. Evolution or education may have given us dispositions to err on the side of caution in perceiving predators, to overestimate our own popularity, and so on.<sup>22</sup> But my thesis is not that belief is completely shielded from mechanisms that tend to make it false; my thesis is that belief is necessarily subject to mechanisms designed to make it true.

In most cases, the latter mechanisms retain some influence, despite interference from the former. Arguments can dispel our belief that there is a predator in the shadows, though they may not dispel the visual appearance. Evidence can undermine our belief in our own popularity, though perhaps not an egotistical phantasy to the same effect. Unlike an optical illusion or a phantasy, a biased belief usually responds to indications of the truth, however imperfectly.

Even when a belief is prevented from responding to corrective influences, the fact remains that its regulative mechanisms are being prevented from doing what they were designed to do. A phantasy and a biased belief are alike in that

<sup>21</sup> If belief can be aimed at the truth by the subject's cognitive mechanisms, rather than by the subject himself, then Davidson may well be wrong when he says that "someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error—true belief and false belief" ("Thought and Talk," in *Inquiries into Truth and Interpretation* (New York: Oxford Univ. Press, 1984), 170). The subject himself would be required to grasp or understand the possibility of false belief only if he was himself required to aim at avoiding it. If he can have beliefs by virtue of cognitive mechanisms designed to avoid that possibility, then he needn't be able to grasp or understand it.

Paul Boghossian has objected that the resulting account of truth-directedness is disjunctive and therefore seems *ad hoc*. I think, on the one hand, that the account's being disjunctive would not necessarily make it *ad hoc*. There is a marked difference between beliefs that are formed by deliberate acts of judgment and beliefs that are formed sub-personally. An account of belief should not be faulted for marking this difference. On the other hand, I don't think that the resulting account of belief really is disjunctive. Rather, it is an account of a functional state that has multiple realizations. On the difference between disjunctiveness and multiple realizability, see Jerry Fodor's 'Special Sciences: Still Autonomous after All These Years (A Reply to Jaegwon Kim's "Multiple Realization and the Metaphysics of Reduction"),' in *Critical Condition: Polemical Essays on Cognitive Science and the Philosophy of Mind* (Cambridge, Mass.: MIT Press, 1998), 9–24.

<sup>22</sup> Relevant here is Jerry Fodor's, 'Is Science Biologically Possible? Comments on Some Arguments of Patricia Churchland and of Alvin Plantinga,' in *Critical Condition*, 189–202. Note that some apparent cases of adaptively biased belief may not be cases of belief at all. For example, one may benefit from a bias toward the view that a prospective opponent in iterated prisoner's dilemmas will play the strategy of tit-for-tat. But one needn't believe that he will; one need only assume it. (I owe this point to Michael Bratman.)

they fail to track the truth; but the phantasy has no tendency to track the truth at all, whereas a biased belief is diverted from the truth; and something can be diverted from the truth only against the background of a tendency to track it. To say that belief aims at the truth isn't to say that it can never be misled; on the contrary, it's to say that belief *can* be misled: what can't be misled are phantasies.

### Can't Belief be Defined by its Motivational Role?

Modest as this claim may be, it is ambitious enough to rule out a purely motivational conception of belief. According to the latter conception, all that's necessary for an attitude to qualify as a belief is that it dispose the subject to behave in ways that would promote the satisfaction of his desires if its content were true. An attitude's tendency to cause behavioral output is thus conceived as sufficient to make it a belief.<sup>23</sup> According to my conception, however, an attitude doesn't qualify as a belief unless it also has a tendency to be constrained by input in ways designed to ensure that it is true. I thus conceive of belief as constituted both by its power to cause behavioral output and by its responsiveness to epistemic input.

The claim that belief cannot be characterized solely by its motivational role is similar in form to my earlier claim that it cannot be characterized solely as a matter of acceptance, or regarding-as-true: both claims point to truth-directedness as necessary to complete the characterization. And this similarity suggests a solution to my earlier problem of explaining the nature of acceptance. For in the motivational role of belief we find a plausible account of what it is to regard a proposition as true. We can say that belief involves regarding a proposition as true in the sense that it involves a disposition to behave as would be desirable if the proposition *were* true, by doing things that would promote the satisfaction of one's desires in that case. We can thus interpret the locution 'regarding as . . .' to mean "representing in a way that disposes one to behave *as would be desirable if . . .*"—or just ". . . as if . . .," for short.<sup>24</sup>

<sup>23</sup> A classic exposition of this view is R. B. Braithwaite's 'The Nature of Believing,' *Proceedings of the Aristotelian Society* 33 (1932–33) 129–46, reprinted in *Knowledge and Belief*, ed. A. Phillips Griffiths (Oxford: Oxford Univ. Press, 1967), 28–40. See also D. M. Armstrong, *Belief, Truth, and Knowledge* (Cambridge: Cambridge Univ. Press, 1973), 3–6; W. V. Quine and J. S. Ullian, *The Web of Belief* (New York: Random House, 1978), 9–19; Robert C. Stalnaker, *Inquiry* (Cambridge, Mass.: MIT Press, 1984), 4; Lynn Rudder Baker, *Explaining Attitudes: a Practical Approach to the Mind* (Cambridge: Cambridge Univ. Press, 1995), 154 ff; Daniel C. Dennett, 'Do Animals Have Beliefs?' in *Brainchildren: Essays on Designing Minds* (Cambridge, Mass.: MIT Press, 1998), 323–31.

<sup>24</sup> On this sense of "behaving as if . . .," see Braithwaite, 31–32. Ruth Millikan has suggested to me that the output dispositions relevant to the characterization of a cognitive state are, not its

Under this interpretation, saying that belief cannot be characterized solely by its motivational role is just the same as saying that it cannot be characterized solely as a matter of acceptance, since its being a matter of acceptance just consists in its having that role. But then the same reasons should apply to both claims. The reason why belief cannot be characterized solely as a matter of acceptance, I said, is that acceptance is also involved in other attitudes, such as hypothesizing and imagining. Having identified acceptance with the motivational role of belief, I am now committed to saying that the other attitudes sharing the element of acceptance must also share the motivational role. But does imagining that  $p$ , for example, typically dispose the subject to behave as would be desirable if  $p$  were true?

Well, it does in at least one context: the context of child's play, in which imagining disposes the child to pretend. When a child imagines that he is a nurse, for example, he is disposed to behave as would be desirable if he were a nurse; when he imagines that he is an elephant, he is disposed to behave as if he were an elephant; and so on.

Here I depart from the received version of folk psychology, which frames all motivational explanations in terms of desires and beliefs. A desire-belief explanation of pretending would go something like this: the child wants to behave like an elephant, he knows how to behave like an elephant—or, at least, he knows some behaviors that are recognized as elephant-like under the conventions of the nursery—and he is consequently moved to engage in those behaviors. I reject this explanation, in favor of one that invokes the motivational force of the imagination.

In order to defend this departure, I will have to turn from discussing how belief aims at the truth to a discussion of how the imagination motivates. The one discussion depends on the other, however, since the necessity of truth-directedness to the characterization of belief depends on the insufficiency of a purely motivational characterization. I'll return to truth-directedness shortly.

*Behavior motivated by imagining: make-believe*

As I see it, the desire-belief explanation of pretending makes the child out to be depressingly unchildlike. According to this explanation, the child keeps a firm grip on reality while mounting an appearance conceived as such. He puts on an act—an elephant-act—conceived by him as a means of impersonating

actual dispositions, but rather the dispositions that it was designed to have. Just as belief need only be regulated in ways *designed* to track the truth, so it need only be *designed* to produce behavior that would be desirable if the belief were true. Whether it actually tends to produce that behavior is no more relevant to its being a belief than whether it actually tends to track the truth. I like this suggestion, and yet I'm not sure whether to adopt it. My own linguistic intuitions inclined me to question whether a state without the appropriate motivational force would count as a belief.

something that he is not. His subsequent pretending is a case of purposeful simulation, no different from an adult's pretense.<sup>25</sup>

I call this explanation depressing because it denies that the child ever enters into the fiction of being something other than he is. In order to enter *into* the fiction, the child would have to act it *out*; and in order to act it out, I think, he would have to act *out of* imagining it, not out of a desire to represent it in action.<sup>26</sup> A child who was motivated by such a desire would remain securely outside the fiction, thinking about it as such—that is, as a fiction to be enacted. I'll expand on this claim below; but first, let me mention some further drawbacks of the desire-belief explanation of pretending.

One further drawback is that the desire-belief explanation fails to account for children's ability to invent and to understand novel ways of pretending. An especially imaginative child may come up with his own way of pretending to be an elephant, but not by considering which behaviors would be most suitable to an elephant-act, as if he were an impressionist honing some zoological schtick. Rather, the child's method is to imagine being an elephant—weighing a ton, walking on stumpy legs, carrying floppy ears—and then to wait and see how he is disposed to behave.<sup>27</sup>

Similarly, this child's playmates do not appreciate his inventions by recognizing that they are especially similar to the behavior of real elephants, and hence good choices for an aspiring elephant-impersonator. On the contrary, success at pretending to be an elephant need not involve behavior that is realistically elephant-like at all. What it requires is rather behavior that's expressive of elephant-mindedness—expressive, that is, of vividly imagining that one is an elephant.

Finally, I think that the present explanation is wrong developmentally, in that it credits the child with a precocious mastery of the distinction between fact and fiction. According to this explanation, pretending entails deliberately producing a false appearance. This explanation should lead us to expect adults to be even better than children at playing pretend: adults are better than

<sup>25</sup> I am going to argue that pretending is not purposeful simulation, but I do not mean to deny that one can enter into a game of pretend with a purpose, or that one may have a purpose in allowing oneself to continue the game. Getting oneself to pretend, and letting oneself go on pretending, may be things that one does for a purpose. But the pretending itself consists in behavior that isn't purposeful, I shall argue, because the particular things that one does in the course of pretending are motivated by wish and imagination rather than by desire and belief. (See the discussion on pp. 272–4 ff. below.)

<sup>26</sup> I am focussing here on one sense in which the child can be said to “enter into” his imaginary world. Walton discusses a broad range of senses in his chs. 6 and 7.

<sup>27</sup> Ian Rumfit has suggested Peter Schaffer's play *Black Comedy* as an illustration of this point. During a large portion of this play, the actors must pretend to be in a pitch black room, though they are in fact playing on a fully lit stage. How is an actor in this play to approach his role? Hardly by mimicking what he has seen people do in the dark. The only answer seems to be: by imagining that he's in the dark.

children at acting—and lying—so why are they worse at pretending? The answer, I think, is that pretending isn't a matter of deliberately producing a false appearance; it's a matter of expressing one's imagination, an activity to which mastery of the distinction between fact and fiction is actually a hindrance, not a help.

I have said that a child's imagining that he is an elephant disposes him to behave as would be desirable if he were an elephant. Yet if imagining is to have a motivational force like that of belief, then it ought to work in concert with an attitude that has a motivational force like that of desire. The behavior to which someone is disposed by accepting *p* is behavior that, if *p* were true, would be desirable in the sense of promoting the satisfaction of his actual desires. Where does desire fit in to my account of pretending?

Note, to begin with, that imagining myself to be an elephant entails other, more specific bits of imagining: here is my trunk, there is my tail, and this—say, this chair—is a tree, or my elephant baby, or a pail of water, or whatever.<sup>28</sup> Within the fiction of being an elephant, I then have various desires with respect to my elephant-world—desires to rub my head against the tree, to drink out of the pail, to nurse my baby, or whatever. And my fictional behavior then expresses my fictional desires: I elephant-do as I elephant-want.

The question is what conation in the real world motivates the behavior that is fictionally motivated by these desires. What conative attitude in me, the child, makes me behave as if I'm an elephant who wants to drink from a pail with his trunk? The desire-belief explanation would be that the actual motive is a desire to simulate the fictional motive—in this case, a desire to act as if I'm an elephant who wants to drink from a pail. According to this explanation, I have no conations with respect to the fictional pail; I have only conations with

<sup>28</sup> The chair then becomes what Walton calls a "prop" (35 ff.). My discussion of this case seems to be at odds with the discussion of imagining that I am Napoleon in my paper 'Self to Self,' *Philosophical Review* 105 (1996) 39–76. Here I say that imagining that I am an elephant entails imagining, of my actual body, that it is an elephant's—that my arm is an elephant's trunk, and so on. In 'Self to Self' I argued that imagining that I am Napoleon need not involve imagining anything about my actual self, since it may involve only imagining the world as experienced by Napoleon.

These arguments are consistent, however. My claim in 'Self to Self' was not that I *cannot* imagine my actual self to be Napoleon; it was rather that I *need* not imagine this in order to imagine that I am Napoleon. There is a way of imagining that I am Napoleon such that the 'I' in this specification of what I'm imagining doesn't refer to me, the imager. My claim in the present paper is not that I must attribute elephanthood to my actual self in order to imagine that I am an elephant; it's rather that I must do so if my imagining is to subserve pretending or make-believe. The reason is that my imagining must refer to my actual body in order to motivate bodily behavior. If I imagined that I was an elephant in the manner described in 'Self to Self,' my imagining wouldn't have the content that I, David Velleman, was an elephant; and so it shouldn't move me to behave as if that content were true—which is what it would do if it moved me to behave like an elephant. (I am grateful to Lucy O'Brien for raising this issue in her commentary on an earlier draft of this chapter.)

respect to the fiction as such, that it *be* a fiction in which there is a conation with respect to the pail.

As noted above, this explanation excludes me from my own imaginary world. It explains how I can be motivated to put things *into* my fiction, including fictional motives, but not how I can act *out* my fiction, in the sense of acting out of my fictional motives. I can't act out of my fictional motives at all, according to this explanation, because they are merely fictional: they don't exist.

An explanation that admitted me into my fictional world would have to allow me real motives toward objects and events in that world: it would have to allow me elephant-desires and elephant-beliefs. In such an explanation, my motivating cognition would not be the thought "Here is how to behave as if this chair were a pail of water"; my motivating cognition would be the thought "Here is a pail of water." And my motivating conation would be, not "Let me behave as if I wanted to drink," but rather "Let me drink." Only by acting out of such motives, framed from the elephant's point-of-view, could I enter into the fiction of being an elephant. So long as I acted out of motives framed from the child's point-of-view, I would remain on the outside of the fiction, looking in.

Consider again the motivating thought "Here is a pail of water," described above as an elephant-belief. This thought is not actually a belief: I cannot believe something to be a pail of water if I know that it is a chair. But the cognition can properly be described as an elephant-belief, because I imagine it to be a belief on my part as an elephant.<sup>29</sup>

Part of what I imagine in thinking "I am an elephant" is that I am an elephant reflecting on what he is rather than a child imagining himself to be what he isn't. Similarly, part of what I imagine with "Here is a pail of water" is that I am an elephant recognizing a pail of water rather than a child re-imagining a chair. I thus imagine my thought "Here is a pail of water" to be a belief.

My imagining this thought to be a belief helps explain how its motivational force enables me to enter into the fiction. When my imagining "Here is a pail of water" moves me to behave toward the chair as would be desirable toward a pail of water, it operates as it would if it were a belief, as I imagine it to be. I therefore act out of motives like the ones that I imagine myself to have. That's why acting out of my imaginings is a way of entering into them: I am motivated as if from within the point-of-view that I imagine occupying.

<sup>29</sup> Walton discusses this phenomenon at pp. 34 and 214–15. Imagining that one (hereby) believes something is analogous in many respects to pretending that one is (hereby) asserting it. On the latter phenomenon, see Walton, pp. 220–24. Also relevant here is the case in which a movie-goer experiences fear-like responses and imagines them to be real fear (241–49). In this case, Walton says that the viewer's responses become props in his imagining (247). Similarly, we might say of a participant in a game of make-believe that his imagining itself becomes a prop, because it is imagined to be a belief.



This phenomenon is summed up by the description of my activity as *make-believe*. I cannot *make* myself *believe* that a chair is a pail of water. But I can imagine that a chair is a pail of water, and that I am thereby believing it to be one; and my imagining can then motivationally simulate its own imagined role, so that it functions as a *mock-belief*. In the motivational grip of this mock-belief, I am gripped as if by the belief that I imagine having as an elephant, and so I am inside the fiction, acting it out.<sup>30</sup>

The previous question, about the role of conations, can now be rephrased as follows. If my imagining "Here is a pail of water" serves as a mock-belief, is there something that serves as the corresponding mock-desire? Could there be an actual state of mind that I imagine to be the desire "Let me drink"?

What serves as this mock-desire could not actually be a desire, for reasons similar to those for which my elephant-belief cannot actually be a belief. Just as I cannot believe something to be a pail of water if I know that it is a chair; so, too, I cannot desire to drink from what I know to be a chair. I usually cannot desire things that are patently unattainable, any more than I can believe things that are patently false. If I think that something cannot come about through efforts of mine, then the most I can do is hope for it; and if I think that it cannot come about at all, then the most I can do with respect to it is wish.<sup>31</sup>

In the present case, what is imagined to be an elephant's desire for an attainable drink must in reality be a child's conation toward a drink known to be unattainable. And if I am to have a conation toward an admittedly unattainable drink, then it must be a wish rather than a desire, just as my cognition of the drink must be an instance of imagining rather than belief. What is imagined to be my elephant-desire for a drink must therefore be a wish. Of course, it isn't an earnest or heartfelt wish; it's a faint and ephemeral wish of the sort that we might ordinarily call a whim. Imagining that I am an elephant, I am struck with the whim of taking a drink.<sup>32</sup>

A motivational explanation of make-believe has now emerged along the following lines. What moves me to dangle my arm between my nose and the seat of a chair is, on the one hand, imagining that this is the way to drink from a

<sup>30</sup> As David Hills has pointed out to me, another respect in which imagining functions as a mock-belief is suggested by Walton's thesis that imagining aims at the fictional. As the child finds his imaginings dictated by what is fictional in the context of the game, he imagines that they are beliefs being dictated by the facts around him.

<sup>31</sup> Of course, I may occasionally find myself desiring patently unattainable things, just as I sometimes find myself believing patent falsehoods; but such cases are exceptions to a rule that I cannot simply and straightforwardly break.

<sup>32</sup> One might think that I need only imagine wanting to take a drink, without necessarily wishing to take one, or even having a whim to that effect. But I do not see how I can be moved by a merely imagined conation. I do think, however, that a whim is the sort of conation that is easy to acquire. Indeed, a whim may be possible to *conjure up*—to form simply by imagining having the corresponding desire.

pail of water with my trunk; and, on the other hand, wishing to drink from a pail of water with my trunk. In the fiction that I am an elephant, my imagining and wishing are a belief and desire, moving me to drink from the pail. When my imagining and wishing move me to behave as if drinking, they fulfill the motivational role of the belief and desire that they are imagined to be, with the result that I enact my imagined role as an elephant.<sup>33</sup>

This motivational explanation of make-believe implies that the attitudes of imagining that *p* and believing that *p* are alike in disposing the subject to behave in ways that would satisfy his conations if they were true. It therefore supports my thesis that the motivational role of belief is not sufficient to distinguish it from other cognitive attitudes. To represent a proposition in a way that confers a disposition to behave as if it were true is simply to regard the proposition as true—which is to have a cognition of it, but not necessarily a belief.

One might think that my explanation of make-believe already suggests

<sup>33</sup> An anonymous referee has proposed an alternative account of make-believe. The referee proposes that imagining that *p* consists in making it the case that it is for oneself as if *p*. Pretending that *p* is a way of making it for oneself as if *p*. Hence pretending is itself a form of imagining, as the referee conceives it. This conception of imagining implies that the child's pretending to be an elephant constitutes—and hence cannot result from—his imagining that he is an elephant.

According to this conception, the child's imagining takes place, not in the medium of mental imagery or a language of thought, but in the outward medium of symbolic behavior, which is not to be explained by any antecedent, inner imagining. The problem with this conception is that, whereas the child has an innate ability to represent things in mental imagery or the language of thought, he must choose his behavioral symbols, and his choices would seem to require some psychological explanation. Why does the child use his arm to represent an elephant's trunk? Why doesn't he use his nose instead? And what makes it the case that he's using his arm to represent the elephant's trunk rather than its tusk, or nothing at all? Surely, the explanation is that the child *thinks* of his arm as somehow corresponding to a trunk. And now the question is how to characterize this thought. The answer cannot be that the child's outward movement plays the role of this thought, too—that he "thinks" of his arm as a trunk just by using it to represent a trunk. This answer would raise the same questions all over again. What would explain the child's "thinking" with his arm rather than his nose? And what would make it the case he was "thinking" of a trunk—or, indeed, that he was "thinking" at all, rather than just waving his arm around? The child's symbolic use of his arm must be explained by some thought lying behind it, and the question is how to characterize that thought. My answer is that the thought consists in the child's imagining his arm to be a trunk. The only alternative I can think of is to say that the thought is a belief, to the effect that his arm is a means of representing a trunk. But then the proposed explanation of make-believe would collapse into the belief-desire account that I criticize in the text.

A closely related problem is this. How does a particular way of moving his arm succeed in making it for the child as if he were an elephant? Does it somehow seem elephant-like to him? If so, then make-believe requires the child to find behavior that strikes him as elephant-like, and it therefore amounts to an elephant-impersonation that the child directs at himself. As I argue in the text, this conception of make-believe portrays the child as implausibly sophisticated and calculating. In my view, a particular way of moving succeeds in making it for the child as if he were an elephant because *it is the way he finds himself disposed to move when he imagines being an elephant*. This account has the virtue of not requiring the child to have any prior conception of how an elephant moves.

a simpler way of distinguishing belief from imagination. The explanation suggests that belief is the attitude that motivates in conjunction with desire, whereas imagining is the attitude that motivates in conjunction with wish. Perhaps, then, a belief disposes the subject to do what would satisfy his desires if it were true, whereas imagining disposes him to do what would satisfy his wishes.

A problem with this suggestion is that belief and imagining may not be so exclusive about the conations with which they combine to motivate behavior. I shall later be discussing cases in which imagining seems to motivate in conjunction with desire;<sup>34</sup> and there may also be cases in which belief motivates in conjunction with wish.<sup>35</sup> Furthermore, even if belief motivated solely in conjunction with desire, and imagining solely in conjunction with wish, the problem would remain how to distinguish the one kind of pair from the other. Attitude-pairs of both kinds dispose the subject to behave in ways that would promote the satisfaction of the one attitude under circumstances satisfying the other. Which pairs of such motives are the belief-desire pairs? If we try to fix the motivational characterization of belief by pairing belief and desire, we end up with an inadequate motivational characterization of belief-desire pairs.

My motivational explanation of make-believe gains support from the experience of adults who have difficulty joining wholeheartedly into this activity. Consider the adult who freezes when invited by a child to join the other elephants at the watering-hole. What this adult experiences is an inhibition, but it isn't at bottom an inhibition against the requisite outward behavior. He could force himself to go through the outward motions of participating in the make-believe while still being inhibited from actually participating. The reason is that he could still be inhibited from acting out his imaginings, which is what make-believe requires.

As I suggested earlier, this inhibition is acquired on the way to adulthood, in the process of mastering the distinction between fact and fiction, or between fantasy and reality. I can now explain further how I conceive of that process.

Mastering the distinction between fantasy and reality requires that we learn to seek what we can actually obtain, and to seek it in ways by which we can actually obtain it. It therefore requires that we come to suppress some of our behavioral dispositions. We have to suppress our dispositions toward trying to make thoughts true, if they cannot be made true, or not by us; and we have to

<sup>34</sup> Consider my variation on Hume's case of a man suspended in a cage (pp. 269–72, below). This man doesn't merely wish that he wouldn't fall; he wants not to fall. But he imagines that clinging to his cage will provide additional safety.

<sup>35</sup> Such cases seem harder to come by, because the distinction between desire and wish is less clear. Do you want to win the lottery, or do you merely wish that you would win? Certainly, you believe rather than merely imagine that buying a ticket is the only means of winning. Are you then moved to buy a ticket wishfully or desirously?

suppress our dispositions to behave as would be desirable if thoughts were true, if they aren't actually true. Both sets of dispositions lead to behavior that's unrealistic, either because it has unattainable ends in view or because it adopts ineffective means. We need a way of suppressing our dispositions toward such unrealistic behavior.

Our solution, I think, is first to segregate our realistic conations and cognitions from their unrealistic counterparts, and then to acquire an inhibition against the motivational force of the latter. Among the thoughts that we are disposed to make true—that is, among our conations—we delimit a subset whose members we are disposed to revise, discard, or at least reclassify if we cannot actually make them true. These reality-tested conations are our desires, which interact with one another in relative isolation from our mere hopes and wishes. Similarly, among the thoughts for which we have a disposition to behave as would be desirable if they were true—that is, among our cognitions—we delimit a subset whose members we are disposed to revise, discard, or at least reclassify if they aren't actually true. These reality-tested cognitions are our beliefs, which interact with one another in relative isolation from our mere imaginings. Setting our desires and beliefs apart from our wishes and imaginings is the first step toward mastering the distinction between fact and fiction.

The second step, I suggest, is to develop an inhibition against the motivational force of the unrealistic attitudes.<sup>36</sup> This inhibition tends to prevent us from manifesting the dispositions to make-true that are associated with thoughts unregulated for practicability; and it tends to prevent us from manifesting the dispositions to behave-as-would-be-desirable-if-true that are associated with thoughts unregulated for truth. It tends to prevent us, in other words, from manifesting the motivational force of wishes and imaginings, so that we tend to act only on desires and beliefs. We thus learn to behave realistically, out of conations that have been constrained by what is attainable, and cognitions that have been constrained by what is the case.

#### *Behavior motivated by imagining: talking to oneself*

Our inhibition against being motivated by unrealistic attitudes is perhaps clearest when it is less than fully effective. One such case is the behavior that is ordinarily called talking to ourselves.<sup>37</sup>

<sup>36</sup> I am now using the term 'realistic' to mean "tested against reality." Of course, we sometimes use the term to mean "properly and successfully tested against reality," so as to exclude attitudes that have failed reality-testing, or whose reality-testing has been inadequate. Thus, there is the sense in which wishes and fantasies are unrealistic attitudes, and then there is the sense in which some desires and beliefs are less realistic than others. From here on, I'll be using the term in the former sense.

<sup>37</sup> One of the "arational actions" discussed by Rosalind Hursthouse, in her paper by that title, is "muttering imprecations under one's breath" which can be an instance of talking to oneself in the sense that I have in mind ('Arational Actions,' *Journal of Philosophy* 88 (1991), 64). I am in

By “talking to ourselves,” I don’t mean literally addressing remarks to ourselves, as when we give ourselves a reminder or a scolding. The case that I have in mind is the one in which, though described as talking to ourselves, we are actually imagining ourselves in conversation with someone else, saying things that we wish we had said or could say. We walk down the street muttering at an invisible interlocutor, perhaps even shaking our head for emphasis; we sit at a red light and tick off our points on the steering wheel; or we sit at the computer alternately writing a sentence and reading it to an imagined audience.

This behavior eludes desire-belief explanation.<sup>38</sup> There is nothing that we both want to do and believe ourselves to be doing by talking to ourselves in this way. If someone stopped us on the street and asked “Why were you just muttering and shaking your head like that?” we could not offer an answer that began with the words “I wanted. . . .” What could we have wanted? To walk along muttering and shaking our heads? Hardly.

An explanation of our behavior would have to begin “I imagined . . .” or “I wished . . .” Indeed, the explanation would have to include both. Merely wishing to say something to someone wouldn’t have made us move our lips unless we had imagined ourselves in conversation with him. And merely imagining ourselves in conversation with someone wouldn’t have made us move our lips if there hadn’t been something that we wished to say.

agreement with Hursthouse’s negative thesis about such actions—namely, that they cannot be explained as motivated by desire and belief. But Hursthouse thinks that they have no further explanation than that the agent, in the grip of an emotion, felt like performing them; whereas I believe that they can be explained as motivated by the imagination.

<sup>38</sup> Desire and belief can move us to initiate an imaginary conversation—for example, when we want to prepare for a real conversation in the future. But what desire and belief move us to initiate in that case is an adult form of make-believe (or role-playing, as we might call it), which then proceeds under the guidance of wish and imagination.

Thus, at the point in the imaginary conversation when we retort with “So’s your old man,” we aren’t moved by a desire to prepare for a future utterance of this retort; we’re moved by the wish to insult our imagined interlocutor. Our only reason for wanting to rehearse such a remark in the imaginary conversation would be that we foresaw wanting to deliver it in the real one; but our only way to foresee wanting to deliver it in the real conversation would be by finding ourselves moved to deliver it in the imaginary one; and at that point, we would already have rehearsed it, thus preempting any desire to do so. That’s why role-playing, even if initiated by desire and belief, must still consist in behavior motivated by wish and imagination. Even when we talk to ourselves because we want to prepare for a future conversation, we have to say things that we wish we were saying to our prospective interlocutor, rather than things that we want to say to ourselves. (This argument bears some similarity to that at the end of n. 33, above.)

In any case, deliberately imagined conversations are not what I am discussing at this point in the text. I am discussing conversations that we have no reason for wanting to imagine, because they are already over or could never take place. Our only motive for imagining these conversations, to begin with, is wishing that we had or could say something that we didn’t or can’t. (The attribution of wishes in this note is subject to difficulties that I discuss in nn. 40 and 43, below.)

The normal inhibition against acting on such unrealistic motives slackens when we talk to ourselves, but it doesn’t entirely let go. It is still evident in the fact that we talk under our breaths. If we entirely lost this inhibition, we would address our imagined interlocutors right out loud, and the sidewalks would be filled with a babble of half-conversations. The fact that we talk to ourselves under our breaths suggests that the inhibition against unrealistic motivation is selective: it prevents behavior that would be inconvenient or self-destructive, but it permits behavior that is harmless, despite being unrealistic.

Make-believe and talking to oneself differ in two circumstantial respects. The imagining involved in these activities is anchored differently in the surrounding conditions and cognitions.

When I make believe that I am an elephant, my imagining is realistically motivated by a desire. I want to imagine that I am an elephant, and this desire moves me to imagine that I am an elephant, with the result that I do what I want.<sup>39</sup> My imagining is itself a mental action motivated by a desire for some result and a (trivial) belief about how to attain it. When I talk to myself, however, my imagining is motivated by the same wish that makes me move my lips. What moves me to imagine myself conversing with someone is, not wanting to *imagine* conversing with him, but rather wishing that I *were* conversing with him.<sup>40</sup> Hence my imagining is not a way of doing what I wish; indeed, it is a way of not doing what I wish, by imagining it instead. It is thus an unrealistic, wishful mental activity rather than a realistic, goal-pursuing action.<sup>41</sup>

Another difference between make-believe and talking to myself is that, in the former, my imagining is accompanied by beliefs about the same behavior. I imagine that I am dipping my trunk into a pail of water, but I also believe that I am dangling my arm over a chair. I know what I’m doing, even as I

<sup>39</sup> On deliberate imagining, see Walton, 13–16, and n. 25, above.

<sup>40</sup> Here my discussion encounters a difficulty in the attribution of wishes. (Until now I have avoided this difficulty by avoiding the construction “wishing that . . .” in favor of the construction “wishing to. . . .”) When we characterize a wish with a that-clause, we are sometimes unable to make our characterization as indefinite as the wish itself. “I wish that I were conversing with him” suggests a determinately present-tense wish—the wish to be conversing even now. Yet the wish itself may have no temporal aspect at all; its propositional content may simply be “I converse with him,” representing a conversation without specifying its temporal relation to the present. Unfortunately, English usage does not permit us to say “I wish that I converse with him.” It forces us to put the dependent verb in the subjunctive, where its tense is unavoidably interpreted as attributing a temporal aspect to the wish. (See also n. 43).

<sup>41</sup> In fact, I think that there are two different mechanisms by which imagining can be motivated. On the one hand, imagining that *p* can be directly motivated by the wish that *p*, through the mechanism of wishful thinking. On the other hand, imagining that *p* can be motivated by the wish that *p* combined with a further bit of imagining, to the effect that one is actually bringing *p* about by imagining it. In the latter case, imagining that *p* is itself a bit of make-believe, acting out the phantasy that Freud called “the omnipotence of thought.” I discuss this phantasy further in n. 48, below.

imagine myself doing something else. When talking to myself, however, I may have no beliefs at all about my verbal behavior. If someone really did stop me to ask why I was muttering and shaking my head, I would probably respond "Was I?"—or perhaps even "I was not!" As I imagined speaking with someone, I was completely oblivious to the fact that I was speaking into thin air.

In these two respects, talking to myself resembles some further, less familiar cases of motivation by imagining. I therefore turn to a new category of examples.

*Behavior motivated by imagining: psychoanalytic examples*

Freud draws our attention to a passage in which Goethe recounts his earliest childhood memory, as follows: he was throwing crockery out a window and watching it smash in the street, to the wicked delight of older boys.<sup>42</sup> Freud tells us that he thought nothing of the passage until a patient, who was unacquainted with the works of Goethe, remembered doing precisely the same thing. What was significant for Freud was that the patient dated his memory to the age at which he had also attacked his baby brother in the cradle. Freud hypothesized that in casting out the crockery, the child phantasized that he was casting out his baby brother.<sup>43</sup> His behavior was thus motivated, according to

<sup>42</sup> 'A Childhood Recollection from *Dichtung und Wahrheit*,' *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, ed. James Strachey et al. (London: Hogarth Press, 1958), xvii, 146–56.

<sup>43</sup> I will reserve the term 'fantasy' (and its homonym 'phantasy') for imaginings that are generated by wish rather than purposely formed by the subject, out of a desire to imagine. I use the spelling "fantasy" for those wishful imaginings which are self-consciously unrealistic, in that they are accompanied by an occurrent awareness of their being in tension with the truth. We sometimes allude to this awareness in specifying the fantasy's content. When we say, for example, that someone has a fantasy of being taller than he is, we do not mean that he has a fantasy with the content "I am taller than I am." We usually mean that he has a fantasy of being a particular height, colored by the knowledge that he isn't. For wishful imaginings that are not self-consciously unrealistic in this sense, I use the spelling "phantasy." Thus, a person cannot phantasize being taller than he is, in the sense just explained, because a phantasy (so spelled) would exclude any awareness of the contrast between his imagined and actual height.

The term 'wish', like the term 'fantasy', tends to be interpreted as denoting an attitude that is self-consciously unrealistic—in this case, a conation felt to be in tension with what is practicable. Here again, this awareness tends to be incorporated into our specification of the attitude's content. When we say that someone wishes that he could be six feet tall, we do not mean that the object of his wish is a mere ability, the ability to be six feet tall; we mean that he wishes he were six feet tall but realizes that he isn't and, more importantly, that he can't be. Saying "he wishes that he could be" is our way of highlighting the point at which the subject feels his wish to be in tension with what is practicable. Unfortunately, wishes that are not self-consciously unrealistic in this sense have no term of their own. I am tempted to coin the spelling "whish" for this purpose. Whishes would be unrealistic conations that are not accompanied by any thought as to the impracticability of their object.

The English language is not well equipped for the attribution of wishes. If someone has an unrealistic conation toward the proposition "I am six feet tall," we are forced to use the subjunctive in specifying the content of his attitude: we say "He wishes that he were six feet tall."

Freud, by a wish to be rid of his infant rival.<sup>44</sup> Freud then reads the same interpretation back into the passage quoted from Goethe, with the help of data about the birthdates of the poet's younger siblings.

Elsewhere Freud tells of an obsessive patient who "used to repeat an especially noticeable and senseless obsessive action":<sup>45</sup>

She would run out of her room into another room in the middle of which there was a table. She would straighten the table-cloth on it in a particular manner and ring for the housemaid. The latter had to come up to the table, and the patient would then dismiss her on some indifferent errand. In the attempts to explain this compulsion, it occurred to her that at one place on the table-cloth there was a stain, and that she always arranged the cloth in such a way that the housemaid was bound to see the stain.

Freud explains this bizarre behavior as follows:

The whole scene proved to be a reproduction of an experience in her married life which had later on given her thoughts a problem to solve. On the wedding-night her husband had met with a not unusual mishap. He found himself impotent, and 'many times in the course of the night he came hurrying from his room into hers' to try once more whether he could succeed. In the morning he said that he would feel ashamed in front of the hotel housemaid who made the beds, and he took a bottle of red ink and poured its contents over the sheet; but he did it so clumsily that the red stain came in a place that was very unsuitable for his purpose.

Freud concludes: "With her obsessive action, therefore, she was representing the wedding-night."

The "problem" posed by this patient's wedding night might just have been that her husband was impotent and that she wished he wasn't. But other symptoms attributed to this patient suggest that the problem was more complicated: she couldn't leave her husband without casting doubt on his potency, and she wished that she could. Her solution was to phantasize that the table-cloth was a bed-sheet, and that the stain on it enabled her to prove her husband's potency, as a previous stain had not.

Unfortunately, the subjective "were," which marks the contrast between content and reality, is normally interpreted as attributing an awareness of that contrast to the subject. "He wishes that he were six feet tall" is thus understood to describe a subject whose conation toward being six feet tall is accompanied by the awareness that he isn't and can't be. We have no way of saying that the subject has a conation toward being six feet tall without any thought as to how tall he could actually be.

These remarks on wish and phantasy are intended to be a gloss on some of Freud's views about primary process (as expressed, for example, in part V of 'The Unconscious,' *SE* xiv, 161–215, at 186 ff). For another difficulty in the attribution of wishes, see also n. 40, above. Putting these two notes together, I would be inclined to argue that wishes necessarily lack a temporal aspect, and that this phenomenon is what Freud had in mind in saying that unconscious processes are "timeless" (187).

<sup>44</sup> *Ibid.*, 152.

<sup>45</sup> 'Obsessive Actions and Religious Practices,' *SE* ix, 116–27, 121.

Freud's explanations for these behaviors make them seem somewhat like make-believe, and somewhat like talking to oneself, but also significantly different from both. They consequently provide us with a third set of circumstances in which imagining can motivate.

Freud's hypothesis is not that the jealous child was playing a game of Throwing Out the Baby, or that the dissatisfied wife was playing a game of Show the Stain, though this image of symptomatic behavior as make-believe can perhaps serve as a first approximation to Freudian theory. One difference is that the phantasies that motivate these behaviors are themselves motivated by the associated wishes rather than a desire to phantasize. What leads the child to phantasize that he is throwing out the baby is, not that he wants to *phantasize* doing this, but rather that he wishes to *do* it. Similarly, the wife wishes that she could prove her husband's potency; she doesn't want to phantasize proving it.<sup>46</sup>

In this respect, symptomatic behavior is more like talking to oneself than it is like make-believe. The wife is like someone who wishes that some fateful conversation had gone differently, and who is thereby moved repeatedly to imagine its going differently and to speak her part in the re-imagined conversation under her breath. In this case, she wishes that her honeymoon had gone differently, and she is thereby moved to phantasize its going differently, though she is moved to play her part, as it were, right out loud.

The wife's failure to mute her symptomatic behavior points to a respect in which it differs from talking to herself. In talking to herself, she would usually be aware of the conversation that she was imagining, though perhaps oblivious to moving her lips. In calling the housemaid to the table, however, she is aware of her overt behavior but oblivious to the phantasy that it enacts. Hence the pattern of awareness is reversed, and this reversal may explain why her behavior isn't muted, like speech under the breath. Because she is unaware of acting out a phantasy, she cannot be restrained by any inhibition against unrealistic motivation.<sup>47</sup> She is aware only of calling the housemaid to the table, a behavior against which she has no inhibition.<sup>48</sup>

<sup>46</sup> On the expression "she wishes that she could," see n. 43, above.

<sup>47</sup> As Jerry Neu has pointed out to me, I am oversimplifying the connection between the subject's overt behavior, on the one hand, and her ignorance of its motivation, on the other. I point out that the subject feels free to produce this behavior because she is unaware of the phantasy that it enacts. But there is another connection: she acts out her phantasy symbolically, in behavior that she feels free to produce, partly in order to protect herself from awareness of the phantasy. This latter connection belongs to the mechanisms of repression, which are beyond the scope of the present paper.

<sup>48</sup> My account of how phantasy figures in these instances of Freudian explanation differs from an account favored by other philosophers. (See, e.g., Richard Wollheim, 'Wish Fulfillment,' in *Rational Action*, ed. Ross Harrison (Cambridge: Cambridge Univ. Press, 1979), 47–60; Sebastian Gardner, *Irrationality and the Philosophy of Psychoanalysis* (Cambridge: Cambridge Univ. Press, 1993).) These philosophers emphasize the role of phantasy in placating or gratifying desire. They consequently conceive of phantasy as shorting the motivational circuit that connects conation with outward behavior: conation produces and is placated by phantasy instead. Wollheim's

*Behavior motivated by imagining: expressive behavior*

Another kind of behavior that is often motivated by phantasy is behavior that's expressive of emotion.<sup>49</sup> Hume points out that a person who is suspended at a great height in a metal cage may tremble with fear despite knowing that he is securely supported.<sup>50</sup> Hume's point is that, although the person doesn't believe that he's going to fall, he does imagine falling, and imagination can arouse the

account of this motivational short-circuit strikes me as question-begging. In order to explain how the imagination gratifies desire, he hypothesizes that it produces belief; and in order to explain how the imagination produces belief, he credits the subject with an underlying belief in the "omnipotence of thoughts"—that is, in his own ability to produce states of affairs by imagining them. But surely, the omnipotence of thoughts is a phantasy. The claim that this product of the imagination is actually believed would itself call for some explanation, which could hardly invoke the subject's belief in the omnipotence of thoughts.

A further question raised by the present interpretation is this. If phantasy shorts the motivational circuit connecting conation with behavior, then why does any behavior ensue? Why does young Goethe actually throw crockery out of the window, if his wish can be gratified internally, by phantasies of expelling his younger sibling? Gardner's answer is that the subject's behavior is expressive rather than instrumental: it doesn't aim at producing wished-for results; it merely expresses the phantasy that has gratified the subject's wish for them. (See pp. 169–72 of Gardner; see also J. Balmuth, 'Psychoanalytic Explanation,' *Mind* 74 (1965) 229–35.)

I think that Gardner is too quick to deny that the behavior in question is aimed at producing results. Of course, it isn't *realistically* aimed at producing them. That is, the wish for some results does not move the subject to do what he *believes* will produce them; but it does move him to do what he *phantasizes* as producing them. When Gardner denies that the subject's behavior is instrumental, he seems to mean that it isn't realistically instrumental, which is true, but he thereby seems to ignore the possibility of its being phantastically instrumental instead. The latter possibility shows that the instrumental and the expressive are not mutually exclusive categories of behavior: behavior often expresses phantasies precisely in the unrealistic way that it aims at producing wished-for results.

Even phantasies themselves can be instrumentally motivated in this fashion; and here the "omnipotence of thoughts" appears in its proper role, as a phantasy rather than a belief. What moves a subject to conjure up wished-for results in his imagination may be the phantasy that he is thereby producing them in reality. Imagining what he wishes for is then a piece of phantastically instrumental behavior, motivated by conation and cognition. The motivating conation is a wish for some results; the motivating cognition is the phantasy that he is producing those results, when he is in fact only imagining them. In such a case, the motivational circuit really is shorted. For if someone imagines that he is producing the wished-for results by imagining them, then he usually does not go on to imagine producing them by means of additional, outward behavior. I suspect, however, that most phantasies are motivated by wishes without any help from the phantasy of omnipotence: they are instances of directly wishful thinking. (See n. 41, above.)

<sup>49</sup> On this topic, see Rosalind Hursthouse, 'Arational Actions.' I share with Hursthouse the view that actions expressive of emotions are not to be explained as motivated by desire and belief. But Hursthouse believes that "[their] only explanation is that, in the grip of the relevant emotion, the agent just felt like doing them" (61). I think that we can often explain why the agent's emotion made him feel like doing these things rather than other things. The explanation involves phantasies that are naturally associated with the emotion's propositional content.

<sup>50</sup> *A Treatise of Human Nature*, ed. L. A. Selby-Bigge and P. H. Nidditch (Oxford: Clarendon Press, 1978), I.iii.13, p. 148. For other cases of emotional responses to what is merely imagined, see Richard Moran, 'The Expression of Feeling in Imagination,' *The Philosophical Review* 103 (1994) 75–106. See also Jerome Neu, 'A Tear is an Intellectual Thing,' *Representations* 19 (1987) 35–61.

same emotions as a belief.<sup>51</sup> Hume might have added that imagination can also motivate the same behavior as belief, since the person in this example may not only tremble but also cling to the bars of his cage, despite the knowledge that he is thereby gaining no additional safety.

When this person is lowered to the ground, he may rattle the bars of his cage in his impatience to get out. Does he believe that he can rattle his way out of the cage? Probably not. But his impatience will just consist in the wish that he could escape from the cage more quickly, and he will be imagining a quicker way out.

Why do you scratch your head when you're puzzled, hold your head when you're worried, or smack your head when you've made a dumb mistake?<sup>52</sup> Are these gestures a kind of sign-language? And then, if no one else is in the room, are you talking sign-language to yourself? No, you're acting out corporeal images of your own thinking—your mind's body-image, so to speak. You're acting out the phantasy of your memory as a balky machine (or a balky child), your curiosity as an itch, or your worries as raising the pressure inside your skull. You wish that you could jar your memory (or punish it), scratch your curiosity, or contain your worries. Your behavior is thus motivated by wish and imagination rather than desire and belief.

Why do you cower in fear, hide your face in shame, clench your fists in anger, shake your head in regret? There are fantasies at work here, fantasies of shrinking, disappearing, fighting someone, or undoing something.<sup>53</sup> These behaviors,

<sup>51</sup> Walton denies that we have real emotions toward merely imagined objects and events; what we have, he argues, are physiological and psychological reactions that we imagine to be emotions. According to Walton, then, we don't fear fictional characters but only imagine fearing them. I favor a somewhat different hypothesis, that our reactions to fictional characters should be understood on the model of mock-desires and mock-beliefs. I am inclined to think that emotions can be reality-tested, and that our terms for particular emotions properly refer only to their realistic instances. Thus, the term 'fear' is reserved for a response that is somehow regulated in a way designed to make it correlate with real dangers; but we also experience an unregulated version of the same response, which constitutes a kind of fantasy-fear. Unlike Walton, I do not conceive of fantasy-fear as a response that is merely imagined to be fear; I conceive of it as non-reality-tested fear, experienced in the unrealistic mode that is characteristic of the imagination. What we feel toward fictional characters, then, is a real emotion, but it is the real emotion of fantasy-fear. (These remarks are indebted to—though not ultimately in agreement with—views expressed by Jonathan Lear in 'Restlessness, Phantasy, and the Concept of Mind,' *Open Minded: Working Out the Logic of the Soul* (Cambridge, Mass.: Harvard Univ. Press, 1998), 80–122.)

<sup>52</sup> I owe this example to David Hills, who also suggested looking to Hume for other examples.

<sup>53</sup> Also relevant in this context is Jennifer Church's claim that the emotions are "internalized actions" ('Emotions and the Internalization of Actions,' (MS) (published in French as 'L'Emotion et L'interiorisation des actions,' in *La couleur des pensees*, s. dir. P. Paperman & R. Ogien (Paris: Editions de l'Ecole des Hautes Etudes en Sciences Sociales, 1995), 219–36).

Stuart Hampshire proposes a different explanation for expressive behaviors. When you clench your fists in anger, according to Hampshire, you are beginning and then cutting short the aggressive behavior to which anger necessarily disposes you. Similarly, "the man who covers or shrinks, only sketching the action of flight, makes a gesture, or assumes a posture, that is the suggestion

like many of the others mentioned in this section, have an expressive or communicative role, but they are not motivated by desires to express or communicate anything: when you're afraid, you don't cower out of a desire to express or communicate your fear.<sup>54</sup> Rather, I would say, the expressive or communicative role of these behaviors is what wins them some reprieve from the normal inhibition on unrealistic motivation. As an adult, you allow yourself to act out your fantasies insofar as doing so is expressive of your emotions.

Acting out fantasies may go beyond mere gestures. Consider, for example, why setting your watch ahead by a few minutes helps to prevent you from being late.<sup>55</sup> You know that your watch is a few minutes fast, and so looking at your watch never leads you to believe that you are a few minutes later than you actually are. What, then, makes you hurry? Surely, the answer is that you are motivated by a cognition other than a belief.

When you're in the passenger seat of a speeding car, why do you press your foot to the floor? When you're behind the wheel, why do you yell at the drivers of the cars in front of you? And why do you yell at the referees of a sporting event that you're watching on television? I know of no satisfactory explanation of these behaviors in terms of desires and beliefs. I can of course concoct desire-belief explanations for them: desire-belief explanations are all too easy to concoct. But the resulting explanations aren't satisfactory, because they make your behavior look realistically purposive, when it is in fact utterly fantastic.

#### *Motivational differences between imagining and belief*

I have now introduced several categories of examples that feature motivation by imagining. These examples show that imagining that *p* and believing that *p* are alike in disposing the subject to do what would satisfy his conations if *p* were true, other things being equal. Admittedly, the examples have also

of the action, with the effective remainder of it removed" ('Feeling and Expression,' in *Freedom of Mind and Other Essays* (Princeton: Princeton Univ. Press, 1997), 143–59, at 145–46). Hampshire's hypothesis may be plausible in the case of anger, but it is quite implausible in most other cases. Cowering is not truncated or ineffective flight, nor does it sketch or suggest that action. What it suggests is a perfectly fantastic "action," to which Hampshire himself alludes with a figurative synonym: it suggests the action of *shrinking*. When you cower, you behave as if you could become small and inconspicuous at will; and this behavior would seem to be motivated by fantasy.

Of course, cowering may not have a motivational explanation: it may just be an innate, primitive behavior associated with fear. As Ruth Millikan has pointed out to me, even dogs and cats cower when they're afraid. I'm not sure that what this fact proves, however, since I'm not sure that dogs and cats don't have imaginations. They are certainly capable of *playing*, in a way that is strongly suggestive of make-believe; they are also capable of dreaming. In any case, I concede that these expressive gestures are among the more controversial of the examples that I discuss.

<sup>54</sup> Hursthouse makes the same point on pp. 60–61.

<sup>55</sup> I owe this example to a talk by the economist Robert Frank.

suggested that other things are rarely equal between cases of imagining and believing, and hence that the actual manifestations of these states are often different. But these differences do not undermine my thesis.

After all, belief itself cannot be characterized in terms of the behavior that it actually causes, since most beliefs cause no behavior at all, and the same belief will cause different behavior in different psychological contexts. Belief can be characterized only in terms of its disposition to produce behavior under various conditions, such as the presence of a relevant conation and the requisite motor skills, and the absence of conflicting motives and inhibitions. The examples suggest that imagining can be characterized as having the same conditional disposition as belief; the only differences have to do with the satisfaction of the associated conditions.

Thus, for example, most deliberate imagining is accompanied by countervailing beliefs, embodying the subject's knowledge of the facts that he is imagining to be otherwise, such as his knowledge that an imagined pail of water is really a chair. These beliefs exert their own motivational force, which can be expected to compete with that of the subject's imagination. Ordinary beliefs are not regularly accompanied by countervailing beliefs, and so their motivational force encounters less competition.<sup>56</sup> I have also hypothesized that the motivational force of imagining comes under an inhibition, whose effects can be detected, of example, in the way that we lower our voices when talking to ourselves. Both of these differences make imagining less likely to cause actual behavior.

Yet the conditional disposition to cause behavior is the same, and this disposition is all that figures in the nature of belief. The only essential difference between these states is that believing that  $p$ , unlike imagining that  $p$ , is regulated in ways designed to make it reflect the actual truth-value of  $p$ . That's why truth-directedness is essential to the characterization of belief.

#### *Why motivational accounts seem right*

Motivational characterizations of belief seem to dispense with reference to truth-directedness only because they tacitly restrict themselves to realistic motivation, whose cognitive component is necessarily truth-directed. For example, we might define belief as the state that determines the means by

<sup>56</sup> Paul Boghossian has pointed out to me that countervailing beliefs are present in the child who plays make-believe. Why don't they outweigh the motivational force of the child's imagining? The answer may be that the motives of a young child are less well integrated than those of an adult. A young child tends to act, not on the vector sum of all of his motives combined, but on whichever motive is at the front of his mind. That's why he is so impulsive and has difficulty making stable choices or postponing gratification. The child's ability to lose himself in a game of make-believe may thus be related to his tendency more generally to get lost in the salient motive of the moment.

which the subject is moved to pursue desired ends; and we might think that we had managed to distinguish beliefs from fantasies without alluding to whether they aimed at the truth. In fact, however, our definition would tacitly allude to truth-directedness under cover of other terms.

This definition would distinguish beliefs from fantasies only because the behaviors motivated by fantasies—make-believe, talking to oneself, yelling at the television—don't count as instances of pursuing desired ends. Yet the reason why yelling at the television isn't an instance of pursuing an end is, not that it doesn't have envisioned results, but only that it acts out a fanciful conception of how to produce them, whereas the pursuit of an end is essentially behavior whose conduciveness to its envisioned results is taken seriously. To pursue something is, by definition, to implement a serious rather than fanciful idea of how it can be attained; and so the very concept of pursuit already implies motivation by a truth-directed cognition. Defining belief by its role in the pursuit of desired ends would thus conceal but not eliminate reference to its truth-directedness.

I suspect that all motivational characterizations of belief tacitly rely on truth-directedness in similar fashion. Consider, for another example, the notion that belief can be defined in terms of its role in determining what the subject is willing to bet on.<sup>57</sup> This notion is familiar to students of formal decision theory, which is sometimes interpreted as showing how to deduce a person's relative degree of belief that  $p$  from his preferences among various possible gambles on  $p$  (among other propositions). This interpretation of the theory relies, I suspect, on the assumption that attaching payoffs to the truth or falsity of a proposition will induce the subject to get real about it, so to speak, in the sense that he will now respond with his best attempt at getting its truth-value right.<sup>58</sup> Hence the decision-theoretic characterization of belief conceals an assumption of truth-directedness.

Even as innocent a concept as 'action' can be used to smuggle truth-directedness into motivational characterizations of belief—in discussion among philosophers, at least. Philosophers tend to assume that every action involves the pursuit of some desired end, or that every action is performed for a purpose. This conception of action immediately excludes a behavior like yelling at the referees on television, which the agent does for no purpose. All the agent has in this case is a wish, which he only imagines that he is

<sup>57</sup> Thanks to Nishi Shah for suggesting this case.

<sup>58</sup> Not surprisingly, I doubt whether attaching payoffs to the truth or falsity of a proposition will necessarily have this effect. Money and chance are two subjects on which people tend to have powerful phantasies. Asking someone to wager large sums on the proposition that the next toss of a coin will turn up heads may not be a way of eliciting his degree of belief in that proposition; it may instead be a way of stirring up his fantasies of being punished by Fate or contaminated by lucre.

fulfilling; and a wish doesn't lend his behavior any purpose, because a purpose must be something that's realistically pursued.

If we conceive of action as necessarily having a purpose, then we can define belief by its role in motivating action, but only because our concept of action already requires it to be realistically motivated. Properly understood, then, motivational characterizations of belief tend to confirm rather than refute the thesis that belief aims at the truth.

#### *How motivational accounts mislead*

Ideally, conceiving of action as necessarily having a purpose should lead philosophers to withhold the term from unrealistically motivated behavior. In practice, however, it sometimes leads them to see, or to think they see, realistic purposes where none exist. Consider, for example, Donald Davidson's discussion of a case borrowed from Freud.

Davidson describes the case as follows:<sup>59</sup>

A man walking in a park stumbles on a branch in the path. Thinking the branch may endanger others, he picks it up and throws it in a hedge beside the path. On his way home it occurs to him that the branch may be projecting from the hedge and so still be a threat to unwary walkers. He gets off the tram he is on, returns to the park, and restores the branch to its original position. Here everything the agent does (except stumble on the branch) is done for a reason, a reason in the light of which the corresponding action was reasonable. Given that the man believed the stick was a danger if left on the path, and desired to eliminate the danger, it was reasonable to remove the stick. Given that, on second thought, he believed the stick was a danger in the hedge, it was reasonable to extract the stick from the hedge and replace it on the path. Given that the man wanted to take the stick from the hedge, it was reasonable to dismount from the tram and return to the park. In each case the reasons for the action tell us what the agent saw in his action, they give the intention with which he acted, and they thereby give an explanation of the action. Such an explanation, as I have said, must exist if something a person does is to count as an action at all.

Davidson's explanation of this case contains a telling misstatement. He says that it was reasonable for the man, having removed the stick from the hedge, to return it to the path. Why was this reasonable? The man thought that the stick posed a danger in the hedge, but he had previously thought that it posed a danger in the path as well. Why did he remove the stick from one dangerous position only to place it in another position recognized as equally dangerous? Why didn't he throw it even further out of the way? The answer to this question is not available to Davidson, because it entails interpreting the behavior as unrealistically motivated.

<sup>59</sup> 'Paradoxes of Irrationality,' in *Philosophical Essays on Freud*, ed. Richard Wollheim and James Hopkins (Cambridge: Cambridge Univ. Press, 1982), 280–305, at 292.

Freud introduces this example in a footnote to his case study of the patient known as the Rat Man.<sup>60</sup> The behavior described in the footnote belongs to a different patient, but it closely resembles behavior displayed by the Rat Man himself, which Freud describes as follows:<sup>61</sup>

On the day of [his beloved's] departure he knocked his foot against a stone lying in the road, and was obliged to put it out of the way by the side of the road, because the idea struck him that her carriage would be driving along the same road in a few hours' time and might come to grief against this stone. But a few minutes later it occurred to him that this was absurd, and he was obliged to go back and replace the stone in its original position in the middle of the road.

Note that the Rat Man didn't even have the other patient's pretext for returning the object to its original position, since he didn't think that it posed any danger in the new position to which he had moved it; he simply thought that moving it had been absurd. But why, then, did he go back to replace it? Surely, taking the trouble to undo an absurd action is doubly absurd.<sup>62</sup>

Freud's explanation is that the Rat Man is deeply ambivalent toward his beloved and has difficulty coping with his ambivalence toward her, as with various other figures in his life. He has therefore repressed the hostile component of his ambivalence, remaining conscious only of unalloyed love. But his hostility occasionally breaks through the repression, moving him to obsessive actions, in which he enacts murderous phantasies of which he is completely unaware. In replacing the stone in the road, the Rat Man was unconsciously phantasizing that his beloved's carriage would, after all, come to grief against it.<sup>63</sup> Of course, the Rat Man's phantasy was quite unrealistic, as was the conscious thought that originally prompted him to move the stone. There is no reason to believe of a particular stone that it will upset a particular carriage, and so there is no reason for moving the stone in either direction, whether for the sake of protecting or of harming a future passenger on the road. But the Rat Man wasn't motivated, in either instance, by a belief that the stone would harm his beloved; he was motivated by a phantasy of its harming her, which in one instance was mistaken for a belief, and in the other instance was repressed entirely.

Freud explains the behavior of his other, unidentified patient along the

<sup>60</sup> 'Notes upon a Case of Obsessional Neurosis,' *SE* x: 192, n. 2.

<sup>61</sup> *Ibid.*, 190 (emphasis omitted).

<sup>62</sup> Tamar Gendler has suggested to me that the Rat Man may have wanted to escape responsibility for any mishap caused by the stone, by returning it to where it had been anyway, without his intervention. Of course, such an action would actually absolve him of responsibility only if it succeeded in literally undoing his first action, making it the case that he had never moved the stone at all. And the notion that his second action could change the past in this fashion is of course a phantasy, no less than the phantasy that Freud postulates.

<sup>63</sup> 'Notes upon a Case of Obsessional Neurosis,' *SE* x: 191.



same lines. Returning the branch to the path was, in Freud's view, a "hostile act," which wasn't really due to the public-spirited motives adduced by the agent.<sup>64</sup> This explanation cannot be filled out as Davidson recommends. We cannot say: "Given that he wanted to harm someone, it was reasonable to extract the stick from the hedge and replace it on the path." Even if harming someone had been a reasonable project, leaving a branch in the path would not have been a reasonable way of going about it. The correct interpretation of this behavior is therefore unavailable to Davidson, who insists on interpreting it as the realistic pursuit of an end.

The problem for Davidson's theory doesn't depend on the peculiarly pathological features of these cases, such as repressed hostility or obsessive repetitions. A perfectly normal agent may see a stick or stone in his path, imagine its causing a freak accident, and feel compelled to shift it out of the way, well knowing that it is not really a hazard. If we insist on explaining his action in terms of attitudes that would make it reasonable, we shall end up attributing to him a stronger belief in the occurrence of the envisioned accident than he actually had. He didn't believe that the accident would happen; he just imagined its happening, but his fantasy was enough to motivate him. How often have you felt for your wallet or purse after merely imagining a mugging? How often has a mere fantasy of disaster prompted you to check whether you turned off the stove or fastened your seatbelt?<sup>65</sup>

<sup>64</sup> 'Notes Upon a Case of Obsessional Neurosis,' 192, n. 2.

<sup>65</sup> Experimental psychologists have found that 4- to 6-year-old children, if asked to imagine a monster or bunny in a box, will profess to know that the box is empty but, if subsequently left alone with it, will get up from their seats to look (P. I. Harris et al., 'Monsters, Ghosts and Witches: Testing the limits of the fantasy-reality distinction in young children,' *British Journal of Experimental Psychology* 9 (1991) 105-123). Some interpret these experiments as indicating that the children "believe" in the products of their imaginations. Yet as others have pointed out, similar results have been obtained with adults, who clearly do not believe in what they have imagined (Jacqueline D. Woolley and Henry M. Wellman, 'Origin and Truth: Young Children's Understanding of Imaginary Mental Representations,' *Child Development* 64 (1993) 1-17). In the latter experiments, adults were instructed to prepare two containers of sugar water and to label one of them "Sugar" and the other "Cyanide." The subjects then showed reluctance to drink from the second container, though not from the first, despite knowing that the two were equally harmless. (See P. Rozin and C. Nemeroff, 'The Laws of Sympathetic Magic: A psychological analysis of similarity and contagion,' in J. W. Stigler, R. A. Schweder, and G. Herdt (eds.), *Cultural Psychology: Essays on Comparative Human Development* (Cambridge: Cambridge University Press, 1990).) Woolley and Wellman attribute this result to the "emotionally charged" nature of what the adults were invited to imagine; and they suggest that the same explanation may apply to the corresponding results with children. I would argue that the "emotional charge" in these cases simply ensured that the subjects had a conative attitude toward the imagined objects, without which their cognition of the objects would fail to motivate behavior. (See also Angeline Lillard, 'Making Sense of Pretence,' in *Children's Early Understanding of the Mind: Origins and Development*, ed. Charlie Lewis and Peter Mitchell (Hove: Lawrence Erlbaum, 1994), 211-34, at 220-21. For another relevant experiment, see D. M. Wegner, G. Coulton, and R. Wenzlaff, 'The Transparency of Denial: Briefing in the debriefing paradigm,' *Journal of Personality and Social Psychology* 32 (1985) 338-46.)

What these cases tell us, I think, is that tacitly incorporating truth-directedness into our conception of action can lead us to misinterpret people's behavior. Better to incorporate truth-directedness into our conception of belief, and to recognize that beliefs are not the only cognitions that can motivate.

### Answers to Objections

I initially described my thesis as modest, but my arguments for it have now revealed two respects in which it could be even humbler. Let me explain these possible modifications to the thesis and my reasons for resisting them.

#### Must truth be the aim?

My thesis is that belief is an acceptance regulated in ways designed to ensure that its content is true. All that my primary arguments have shown, however, is that belief cannot be fully characterized in terms of its power to cause behavioral output. Even if the upshot of these arguments is that belief must also be characterized in terms of how it is constrained by input, the question remains whether it is constrained in ways designed to ensure that its content is true. Belief might be an acceptance regulated for empirical adequacy, or merely for consistency with other, similarly regulated acceptances. It would then have a constitutive aim other than the truth.<sup>66</sup>

I am less intent upon fending off this objection than some of the others that I have entertained. If I have managed to show that belief must be characterized by its responsiveness to input as well as its power to produce output, I will consider my arguments to have succeeded in some measure, at least. But I do think that the input constraints definitive of belief are designed to yield beliefs that are true, and I'd like to be able to say why I think so.

I take it to be a conceptual truth that beliefs are correct when true and incorrect when false: false beliefs are necessarily faulty or mistaken.<sup>67</sup> What's more, I don't think that the fault in false beliefs can consist in their tendency to misdirect our behavior, since many beliefs have little or no chance of directing behavior, and even some false beliefs can direct us well enough.<sup>68</sup> False beliefs

<sup>66</sup> John Gibbons has suggested to me that the aim of belief is knowledge. This suggestion subsumes my view. If belief is an acceptance that aims at being knowledge, then it aims at the truth and more—i.e., at truth plus proper justification. But I would like the aim of belief to account for our intuition that even an unjustified true belief is correct or right, whereas Gibbons's view would imply that such a belief was a failure as a belief.

<sup>67</sup> On this point, see A. Phillips Griffiths, 'On Belief,' in *Knowledge and Belief*, 127-43.

<sup>68</sup> Most if not all beliefs have some tendency to cause behavior, of course, since they tend to be expressed in assertions. But causing false assertions counts as a form of misdirection, I suspect,

are faulty in themselves, antecedently to and independently of any untoward practical consequences. In what sense are they faulty? The most plausible answer, I think, begins with the observation that we conceive of beliefs as constitutively regulated by input. Faulty or mistaken beliefs are the ones whose regulation has not succeeded in producing the kind of cognitions that it was designed to produce. The fact that beliefs are conceived to be faulty when false indicates that the regulation conceived to be constitutive of them is regulation for truth. Truth-directedness thus appears to be enshrined in our concept of belief.

Our conceiving of belief as truth-directed doesn't necessarily settle the issue, however. Perhaps we could discover that the attitudes we call beliefs are actually regulated in ways designed to promote something other than their being true. Would we conclude that these attitudes weren't really beliefs, after all? Or would we revise our conception of belief, to reflect its newly discovered aim?

I do not have an answer to this question. But I would like to point out, very briefly, the extreme improbability of its premise, that we might discover belief to be regulated for something other than truth.

I think that introspection argues against this possibility. When we discern a gap between a belief and the truth, the belief immediately becomes unsettled and begins to change. If it persists, we form another belief to close the gap, while reclassifying the recalcitrant cognition as an illusion or a bias. I cannot imagine evidence that would show this reclassification to be a mistake.

Some people claim that their cognitive efforts are aimed at something short of the truth, such as instrumental success or empirical adequacy. Under some interpretations, however, this claim is compatible with my thesis. For example, what people call aiming at empirical adequacy may in fact be only a willingness to settle for it. If a basketball player says "I'm just aiming to win by one point," we don't necessarily assume that he is engaged in point-shaving; we assume that he is aiming to score as many points as possible but willing to settle for any margin of victory. His willingness to settle for a small margin is a second-order aim with respect to his degree of success in the first-order aim of scoring points. I don't deny that we have a similar second-order aim with respect to our cognitive effort: we're willing to settle for empirical adequacy rather than truth. What I deny is that we are engaged in epistemic point-shaving, aiming for empirical adequacy in the first instance.

A basketball player may say, alternatively, that his ultimate aim is to raise his own salary, but we don't therefore assume that his every move on the court is

only because of the role that assertions play in forming the beliefs of others—and hence only because of the faultiness of false beliefs.

directed at that aim. The best way for him to raise his salary is to aim at victory for its own sake, without regard to monetary rewards. Money may be the goal for which he enters the game, but he enters the game by adopting its object as an intrinsic aim. Similarly, we may enter the game of having beliefs on a particular subject because we want our motivating cognitions on that subject to yield successful actions; but success in action does not thereby become the object of the game.

As these analogies suggest, the best way of achieving empirical adequacy or instrumental success may be to aim at the truth instead, just as the best way of raising one's salary may be to aim at winning, and the best way of winning may be to aim for the highest possible score. In each case, the former outcomes are much harder to aim at, and aiming at them can be counter-productive.

*Must each belief have the aim?*

Here is another respect in which my thesis could be more modest. I have claimed that our concept of belief is such that each instance of belief must aim at the truth in its own right. But I have also argued that truth-directed cognitions are psychologically segregated from their non-truth-directed counterparts, whose motivational force is generally inhibited; and this argument raises the possibility of defining belief as a cognition that is psychologically grouped with the truth-directed ones, whether or not it is truth-directed itself. Belief would then be defined as a cognition that is *treated as* aiming at the truth, either by the subject or by his psychological mechanisms, in that it is included among those cognitions whose motivational force is uninhibited because of their truth-directedness.<sup>69</sup>

We are all familiar with the experience of finding one of our own mental states on the wrong side of a psychological boundary, as when we find the residue of a dream mixed in with our experiential memories. Ronald Reagan was often reported to have retailed the plots of movies under the guise of historical anecdotes, as if the fantasies in which he had participated as an actor had become mixed in with his beliefs. Reagan, or his mind, treated these cognitions as if they were truth-directed, and yet they were permanently disconnected from the truth, since he went on repeating them, and thus behaving as

<sup>69</sup> In 'Truth, Reason, and the Regulation of Belief,' Peter Railton suggests that belief aims at the truth in the sense it "takes itself to be correct only if [its] content is true," or "presents itself" as "getting things right" (74). In another paper, he says: "It is part of the *price of admission* to belief as a propositional attitude that one not represent one's attitude as unaccountable to truth" ('On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action,' in *Ethics and Practical Reason*, ed. Garrett Cullity and Berys Gaut (Oxford: Clarendon Press, 1997), 53–79, at 57). These statements are similar to the view that I am currently considering, that an attitude qualifies as a belief partly by virtue of being *treated as* truth-directed.

if they were true, no matter how often he was told that they were false. Did the President believe what he was saying?

I am assuming that the President was not deliberately lying. I'm also assuming that he didn't find grounds, however spurious, for discounting the corrections that were offered to him. Finally, I am assuming that he wasn't engaged in any complex form of self-deception, unconsciously orchestrating the evidence that was allowed to enter his thinking. Rather, I assume that he was just as he seemed, blithely impervious to the facts. And on that assumption, I am inclined to think that he did not believe what he was saying. He may have believed *that* he believed what he was saying; but what he was saying conveyed the content of fantasies on his part rather than beliefs.

My reasons for favoring this description of the case derive partly from introspection. Sometimes I find myself wondering whether I really believe a proposition that I'm about to assert. What I ask myself on such occasions is not whether I have found a representation of the proposition among my beliefs. I *have* found it among my beliefs, witness my disposition to assert it. Of course, I may be asking whether I *ought* to believe the proposition—that is, whether it is true—but in at least some cases my question is descriptive with respect to my attitude, not prescriptive. I'm asking whether I even now believe the proposition, or whether it isn't instead the content of a fantasy or assumption that has fallen in among my beliefs. This question could never arise, however, if an attitude's falling in among my beliefs was sufficient for its being a belief. If my mental classification of the attitude determined its nature, then it couldn't be misclassified.

Furthermore, if an attitude's being treated as a belief were sufficient for its being a belief, then misclassified fantasies would tend to fall under epistemic norms. To describe the attitudes expressed in President Reagan's anecdotes as beliefs would imply that Reagan should have discarded them or revised them so as to conform with historical reality. All he should have done, however, was to reclassify them—to re-shelve them, mentally speaking. The President wasn't so much a bad historian as a sloppy mental housekeeper. If we want to cut him this much normative slack, however, we have to think of his anecdotes as reporting the contents of misplaced fantasies rather than irrational beliefs.

I am therefore inclined to resist the suggestion that the acceptance of a proposition can qualify as a belief merely by virtue of being mentally classified among the acceptances whose motivational force is uninhibited because they are regulated for truth. Whether an acceptance qualifies as a belief depends, I think, on whether it is so regulated in its own right.

I don't pretend that our colloquial use of terms like 'fantasy' and 'belief' will always follow my definition in classifying mixed or borderline cases. I am espe-

cially worried about cases of delusion. Aren't there people who believe that they are Napoleon? (People other than Napoleon, I mean.) Don't such people have a belief that isn't regulated for truth?<sup>70</sup>

I think the answer is that it isn't literally a belief. I suspect that we tend to apply the term 'belief' in a figurative sense to phantasies for which the subject doesn't or cannot have countervailing beliefs. When someone is said to believe that he is Napoleon, he actually has a phantasy to that effect; but on the question of who he is, a phantasy is all he has. He is somehow incapable of reality-tested cognitions of his identity. The phantasy of being Napoleon is thus what he has instead of a belief about his identity; and in this sense it is his belief on the topic, just as a cardboard box on the sidewalk may be his house by virtue of being what he has instead of a house.<sup>71</sup>

If you ask me, however, a cardboard box on the sidewalk isn't really a house. And a phantasy of being Napoleon isn't really a belief.

<sup>70</sup> Like President Reagan's historical anecdotes, someone's claiming to be Napoleon may be subject to many different explanations, which would yield different verdicts as to whether he believed what he was saying. Perhaps a person could get himself to think that he was Napoleon by orchestrating the evidence available to him, or by developing elaborate theories discrediting the counter-evidence. At some level, this person's cognition of being Napoleon might then remain under the control of truth-directed mechanisms, which were being diverted from their goal; and in that case, he would literally have deceived himself, by self-inducing a false belief. In the text, however, I am assuming that the subject is simply impervious to the facts of his real identity, and that his cognition of being Napoleon is therefore a phantasy. For a fascinating discussion of some extreme delusions, see Tony Stone and Andrew W. Young, 'Delusions and Brain Injury: The Philosophy and Psychology of Belief,' *Mind and Language* 12 (1997) 327–64. The explanations favored by Stone and Young are consistent with the thesis that belief aims at the truth. See also David Shapiro, 'The Loss of Reality,' in *Neurotic Styles* (New York: Basic Books, 1965), 48–53.

<sup>71</sup> The usage discussed in this paragraph may also apply to phantasies considered in the context of the Unconscious, where there are no reality-tested cognitions at all. Phantasies may qualify as unconscious beliefs, but not because they are beliefs that are inaccessible to consciousness. They may qualify as unconscious beliefs because they are what the Unconscious has instead of beliefs.