

## To Believe is to Know that You Believe

ERIC MARCUS<sup>†</sup>

### ABSTRACT

Most agree that believing a proposition *normally* or *ideally* results in believing that one believes it, at least if one considers the question of whether one believes it. I defend a much stronger thesis. It is *impossible* to believe without knowledge of one's belief. I argue, roughly, as follows. Believing that *p* entails that one is able to honestly assert that *p*. But anyone who is able to honestly assert that *p* is also able to *just say* – i.e., authoritatively, yet not on the basis of evidence – that she believes that *p*. And anyone who is able to just say that she believes that *p* is able to act in light of the fact that she holds that belief. This ability to act, in turn, constitutes knowledge of the psychological fact. However, without a broader theory of belief to help us make sense of this result, this conclusion will be hard to accept. Why should being in a particular mental state by itself necessitate an awareness of being in that state? I sketch a theory that helps to answer this question: believing is a matter of viewing a proposition as what one ought to believe. I show how this theory explains (and entails) the thesis that to believe is to know that you believe.

### 1. Introduction

Many would agree with the following: Believing a proposition *normally* or *ideally* results in believing that one believes it, at least if one considers the question of whether one believes it. Stronger claims in the vicinity – e.g., that believing necessarily results in believing that one believes, that such second-order beliefs amount to knowledge – garner significantly less sympathy. But I will argue for the boldest variation. It is part of the metaphysical profile of belief that the subject has knowledge of her belief. One can no more believe without knowing one believes than an object can be red without being coloured: it's metaphysically impossible. And I'll give the outlines of an account of belief that explains why.

I proceed in two stages. In section 2, I defend an argument that believing is essentially tied to a distinctive kind of knowledge that one believes. Roughly: Believing that *p* entails that one is able to honestly assert that *p*. But anyone who is able to honestly assert that *p* is also able to *just say* – i.e., authoritatively, yet not on the basis of evidence – that she believes that *p*. And anyone who is able to just say that she believes that *p* is able to act in light of the fact that she holds that belief. This ability to act, in turn, constitutes knowledge of the psychological fact. To believe *is* to know, not on the basis of observation or evidence, that you believe.

---

<sup>†</sup> Department of Philosophy, Auburn University, Auburn, Alabama, USA; Email: [marcuea@auburn.edu](mailto:marcuea@auburn.edu)

The argument of the second section, if successful, shows that any theory of belief that fails to recognize its essential self-consciousness is to that extent defective. The argument does not rely on any specific theory of belief. But without a theory of belief, the argument is not credible. For it is quite puzzling why simply being in a particular mental state should by itself necessitate an awareness of being in that state. In the third section, I sketch a theory that helps to resolve this puzzle. Putting caveats and clarifications aside for the moment, I argue that belief is a matter of viewing a proposition as meeting a certain standard of correctness, as being such that one ought to believe it. The ability to say authoritatively, yet not on the basis of observation or evidence, that one *does* believe *p* requires nothing beyond the ability to express one's view that *p* is what one *ought* to believe. Since there is no puzzle about the latter ability, there should be no puzzle about the former.

The two sections are mutually supporting. Section 2 argues, in a theory-neutral way, for a thesis that is hard to accept without an explanation of how it could possibly be true. Section 3 provides a theory that explains (and entails) the thesis, thereby supporting section 2. I advance several independent reasons for adopting the theory of belief sketched in section 3, but one such reason is that it can accommodate the truth about belief established by the argument in the first. Section 2 in that sense supports section 3.

## 2. *Doxastic self-knowledge*

Belief, I argue, is a sort of attitude toward a proposition that includes an awareness of holding that attitude. I defend the thesis by arguing for the following principles:

- (A) S believes that *p* only if S is able to honestly assert that *p*,
- (B) S is able to honestly assert that *p* only if S is able to avow the belief that *p*
- (C) S is able to avow the belief that *p* only if S knows that she believes that *p*.

My argument for these principles establishes that it is impossible that one could believe that *p* yet not know that one believes that *p*. Believing that *p* is metaphysically sufficient for knowing that one believes that *p*. In this section, I will not argue for this conclusion on the basis of a theory of belief. Rather, I'll argue that (A) (B), and (C) are, properly understood, compelling and that there are plausible responses to the obvious objections and counterexamples.

Before considering (A), I'll begin with a methodological preamble. Here is a common strategy for adjudicating philosophical disputes: first, identify the data that need to be explained; and second, consider how well the views on offer explain them. One might expect a defence of my thesis to proceed along these lines. This would be to argue that a theory of belief according to which we are necessarily aware of our beliefs better explains the data than other theories.

But this strategy runs into trouble if the data themselves are under dispute – as they often are, and as they will be here. One controversial implication of the view I aim to defend is that non-rational animals and very young children do not have beliefs – or, more precisely, that the states that we ascribe to them using terms such as ‘believes’ and ‘thinks’ are not the same as those we ascribe to adult humans using those same terms. While such a view is not unprecedented (see, e.g., Davidson 1982 and Stich 1978), many take it to be a constraint on a theory of belief that it allows that young children, human adults and non-human animals all ‘believe’ in the same sense. Unsurprisingly, I will not take this ‘fact’ to be among the data to be explained. (But I will outline, in section 3, a strategy for understanding how – if not in terms of states shared with adult humans – we should understand the minds of non-rational animals and very young children.)

My topic is the rational state that we attribute to normal adult humans using ‘believe’ and its cognates. (And I shall use these terms exclusively to refer to the rational state in what follows, except where otherwise indicated.) This paper puts the following proposition to the test: It is profitable to consider beliefs so understood as constituting a *sui generis* category, characterized (on the basis of the argument of section 2) in part by the inseparability of believing and knowing that one believes. This possibility (and what it would reveal about the nature of belief, were it actual) is invisible if one assumes at the outset that there are non-rational believers, i.e., if one assumes that the state we attribute to rational creatures using the language of belief is also possessed by non-rational creatures. For in that case, knowledge of belief will ‘turn out’ to be only contingently related to belief.

It is no better to leave the issue of whether there are non-rational believers as an open question as we measure rival theories according to how well each explains the data my opponents and I *do* agree about. For an approach that does not specifically attempt to explain (what I argue are) the distinctive epistemic features of belief (understood as a rational state) will not yield an account that illuminates such features. And we shall see in section 3 that what explains these epistemic features is precisely what explains belief’s being rational. Bracketing the question of whether adult humans and non-human animals ‘believe’ in the same sense might seem to be a judicious compromise between assuming that they do and assuming that they don’t. But this methodology simply removes from consideration the idea that the nature of *our* beliefs is essentially bound up with our being rational; it is precisely this idea that I aim to explore here.

It will help to offer a preliminary characterization of what I mean in describing belief as a *rational* state. Belief is an attitude about which one might intelligibly inquire: ‘Why do you believe it?’. This question asks the believer to provide reasons of a particular sort – not those that show what is to be gained for holding a belief but reasons for thinking the relevant proposition is true. The believer may not have a reason or may have a bad reason. But the question asking for her reasons must have

application (to use Anscombe's phrase). Belief, adequately supported by reasons (and sometimes even without any such support), constitutes knowledge. But whereas some beliefs may not constitute knowledge, beliefs are as such eligible to do so.

To understand belief as that to which the specifically doxastic sort of 'why?' question has application is thus compatible with specific instances of rational states being irrational. Belief is a rational state, notwithstanding that beliefs are sometimes irrational. One must not conflate the distinction between the rational and the irrational (a distinction at the level of specific cases) with the distinction between the rational and the non-rational (a distinction at the level of general kinds).

My justification for excluding the states of non-rational creatures also supports the exclusion of various non-rational states of rational creatures, states that might, along with belief proper, also dispose one "to act and react – not just outwardly but also in one's phenomenology and patterns of cognition – shall we say, belief-that-P-ishly" (Schwitzgebel 2011, 45). My thesis does not concern the 'associative, automatic and arational' states that Gendler calls "aliefs", since the question asking for a reason does not apply to these (Gendler 2008). A visceral fear, say, may dispose me to behave as if I believe that failing to step back from a floor-to-ceiling window will lead to my falling out of the building. But I may nonetheless not believe the corresponding proposition – not even a little bit. For what underlies my behaviour is not the sort of state about which a question concerning *my reasons* for being in it gets a grip. There is no question here of *my evidence* for thinking I'll fall. Even if I actually will fall, this state is still not a candidate for *my knowing* that I'll fall. Belief, in the ideal case, constitutes knowledge. But visceral fears are not even defective exercises of our epistemic powers.

Similarly, implicit biases do not fall within the scope of my thesis. For an implicit bias is not eligible to be held on grounds.<sup>1</sup> And so even where implicit biases are accurate, they nonetheless do not count as knowledge. They are non-rational states, and hence not beliefs in the relevant sense. Note that my claim is not that implicit biases (or visceral fears, for that matter) are *insensitive* to evidence. A non-rational tendency might be *causally* sensitive to evidence; but its being non-rational means that the subject does not have an evidential *basis* for being in the state; for it is, qua non-rational, not that sort of state.<sup>2</sup>

<sup>1</sup> I would thus distinguish between implicit bias and tacit belief. A belief that my sandwich is edible is tacit in the sense that it is not 'running through my mind'. But it nonetheless can be held for reasons – I believe it, say, because it's made of bread, salami and butter, all of which are edible. This distinguishes it from an implicit bias. It is no coincidence that tacit beliefs are rational in this sense, and that they are generally taken (notwithstanding their being tacit) to fall within the scope of first-person-authority. One goal of this paper is to reveal the connection between these two facts.

<sup>2</sup> Cf. Schwitzgebel (2010, 539–541). Although the distinction between a state's being caused by p and a state's being held on the basis of p is uncontroversial, there is much dispute about how precisely it should be understood. I weigh in on this issue in Marcus (2012).

I argue (in section 2) that belief is essentially known to the believer; and I give an account of belief (in section 3) that explains this fact – as well as many other salient facts about belief. The illumination provided by this account will, I hope, vindicate my methodology: specifically, my refusal to include among the data to be explained the ‘fact’ that adult humans, very young children and non-human-animals all *believe* in the same sense.

(A) S believes that p only if S is able to honestly assert that p

According to (A), it is a necessary condition of believing that p that one is able to honestly assert it. I employ ‘able’ (as it figures in (A) – (C)) to refer to a two-way power. To say that S is able to honestly assert that p is to say that S has the power both to honestly assert it and not to honestly assert it (see Kenny 1976, ch. 7 and, more recently, Alvarez 2013). It is *not* to say merely that S is disposed to honestly assert it, in the way that fragile objects are disposed to break, nor (if this is something different) that there is some specifiable circumstances under which, necessarily, S will honestly assert that p.

There are nonetheless some respects in which abilities are similar to dispositions. The possession of an ability is consistent with the absence of an opportunity to manifest it, just as the possession of a disposition is consistent with the non-obtaining of its manifestation-conditions. Hector is able to ace the logic exam, even while the exam remains locked up in the professor’s desk. The circumstances must be favourable for the exercise of an ability. Favourable circumstances will sometimes include having the help of others. A brain surgeon might have the ability to remove a tumour, but in order to actually do it he would require *a lot* of help from medical specialists and nurses (not to mention specialized equipment).

Abilities, like dispositions, can also be masked. Hector is able to ace the logic exam, but can’t exercise this ability until the bees stop swarming around him. There is no bar to masking being internal: Hector is able to ace the logic exam, but can’t exercise this ability with that horrible headache. Internal masks can be psychological (e.g., lack of confidence) or non-psychological (e.g., paralysis). It is perfectly consistent to say, then, that S has the ability to x, but can only exercise it if certain obstacles can be overcome.

Like a disposition, an ability to x provides a minimal explanation of a subject’s x-ing, in at least this sense: there can be no question of how it is possible that a subject who is able to x is x-ing.<sup>3</sup> Or, more carefully: given that S has the ability to x, the only remaining questions about how it is possible that S is x-ing pertain to opportunities and masks. Consider the following questions about dispositions: ‘I know it is fragile, but how can it break?’ and ‘I know it is poisonous, but how can it sicken?’. In each case, the question only makes sense insofar as we take

<sup>3</sup> I discuss this and other features of abilities and dispositions in Marcus (2012, ch. 1).

the questioner to know or suspect or wonder about the absence of an opportunity (the cup sits on a shelf, the poison is in the vial) or the presence of a mask (the cup is protected by packing material, everyone has drunk the antidote). If the opportunity-or-mask interpretation of the question is ruled out, we would have grounds for doubting whether the questioner truly possesses the relevant disposition-concept.

This suggests a test for whether a property *p* is sufficient for a disposition or ability to *x*. If *p* is such that ‘I know *S* possesses *p*, but how can *S* *x*?’ is only reasonably interpretable as a question about opportunities or masks, then *p* is sufficient for the possession of the disposition or ability to *x*. By this test, being heavy is sufficient for the disposition to fall (‘I know it is heavy, but how can it fall?’) and being a chess grand-master is sufficient for the ability to beat my five-year old (‘I know she’s a chess grand-master, but how can she beat my five-year old?’). I shall use this test in arguing for (A). I contend that the question ‘I know that *S* believes that *p*, but how can *S* honestly assert that *p*?’ is only reasonably interpretable as a question about opportunities or masks. Later, I will use this test to argue for (B) as well.

No one could dispute that it is at least *normal* for someone who holds a belief to be able to honestly assert it. Once we have put aside cases of non-rational states, such as visceral fears, which (as explained above) are not in the purview of my thesis, the question of whether we can conceive believing that *p* without such an ability would seem to depend on how best to describe certain abnormal cases. Consider Alfred and William, brothers who have both set out for a career as painters, but with very different results. Alfred has been a great success from the beginning: a full scholarship to the Rhode Island School of Design, designated at 23 as one of 10 young artists to watch by Marina Abramović, multiple pieces in the Whitney Biennial by 27, and by 30 considered a leader of a movement of younger artists. But by 34, he has effected a Pynchon-esque disappearance from public view. It is only after Alfred’s whereabouts have become the topic of frenzied speculation, after spotting him has become something of a sport, that William begins to receive any encouragement for his own work. He finds himself invited to various art-world events from which he had formerly been excluded. He is accepted to a prestigious artists’ residency. Important galleries seem more interested in his work, not to the extent of actually offering to represent him, but at least willing to see his work and give advice. William, we can suppose, believes that this newfound interest is not really in his own work, but rather in his connection to Alfred. Indeed, we might even say he *knows* that it is the desire to reach his brother that lies behind these recent overtures. Nonetheless, he cannot overcome the temptation to take pleasure in them as if they reflected genuine appreciation for his heretofore neglected paintings. He begins to tout these developments to friends as evidence that the success for which he has so long pined is now finally

within reach. He flies into a rage at the slightest hint of scepticism among these ‘so-called friends’. The complex brilliance of his work, he vehemently argues, explains both its initial obscurity and its incipient renown. He is self-deceived.

My thesis raises the odd question of whether William nonetheless has the ability to honestly<sup>4</sup> assert that the recent attention he has been receiving is simply the product of interest in his brother’s whereabouts. The issue is not merely the question of whether he is able to grasp and utter the sentence ‘people are only interested in me for the sake of getting in touch with my brother’, but whether he is able to ingenuously use the sentence to say what he thinks. William’s case makes it tempting to think that the ability to honestly assert a proposition involves both the expressed belief and something else – a higher-order belief, for example. To explain how someone who believes that *p* is able to honestly assert it, one needs to appeal not just to the belief, but also to the person’s believing she holds it.

However, this gets the explanatory facts entirely wrong. That someone who believes that *p* is able to honestly assert it needs no explanation. If you tell me that you know that Alma believes that *p*, but wonder how she can say so, your very puzzling question could only be reasonably interpreted as asking about whether some obstacle stands in the way of communication. And so I would naturally follow-up with a question like: ‘What would stop her?’. In other words, I would have to understand your question as pertaining to opportunities or masks. I could *not* reasonably interpret it as pertaining to a missing ability.

William deliberates about the genuineness of the newfound interest in his art – ‘after all, the greatness of truly original works, as Proust observed, is rarely recognized immediately’ – in bad faith, or at least with pathological disengagement. He puts forward the outcome of this disengaged deliberation as his actual opinion, but does not assert what he truly believes.<sup>5</sup> We can sympathize with the obstacle that prevents him from doing so: the dread associated with having to surrender his newfound feelings of self-worth and to return, with renewed despair, to full consciousness of his underlying misery. But regarding the tricky and seldom-asked question of whether this psychological obstacle destroys or merely masks the relevant ability, the general test for dispositions and abilities formulated above reveals it to be the latter. William cannot, so long as he is self-deceived, honestly

<sup>4</sup> I talk about *honest* rather than *sincere* assertion, since some hold that sincerity depends not on whether the speaker believes the asserted proposition, but rather only on whether she takes herself to believe it (see Moran 2005). The same stance would not be plausible regarding *honest* assertion. One cannot honestly assert that *p* unless one believes that *p*. Thus I would argue that someone who doesn’t believe that *p* can’t *honestly* say that *p*, even if Moran is right that she might *sincerely* say that *p*.

<sup>5</sup> Some resist thinking of the self-deceived as truly believing (let alone knowing) the unpleasant truth (see, e.g., Mele 1987, 2001). I note here that if this were correct, then self-deception would not pose a challenge to my view, since there would then be no question of whether the subject could honestly assert the (believed) unpleasant truth.

make the relevant assertion – not because he lacks the ability, but because the ability is masked.

No doubt one could conjure up a philosophical theory of belief that would say otherwise. But note, first, that we have just seen that such a theory is not required in order to accommodate the idea that the self-deceived cannot (while they remain self-deceived) honestly assert what they truly think. Second, the plausibility of the test and what it highlights suggests that it would be a revisionary theory – i.e., one that requires us to revise our concept of belief. Third, there are, as I will show in section 3, independent grounds for preferring a theory that accords with the ordinary conception of belief, according to which believing is inseparable from the ability to honestly assert.

So far I have considered the problem of self-deception as a threat to (A). But the problem re-emerges in slightly different forms in relation to (B) and (C) as well, so we are far from done with the issue.

There are also reasons for doubting (A) that have nothing to do with self-deception. It might be argued that believing that *p* can't by itself explain how someone is able to honestly assert that *p*, for such an explanation must add that the subject possesses a language with which to assert. However, this is not plausibly separate from what's required for belief itself. Note, first, that belief, unlike perception, is considered a paradigmatic *conceptual* mental state. The issue of whether perception is conceptual or non-conceptual has always been framed in terms of whether it is like belief in requiring the perceiver to possess the concepts corresponding to the content of perception. So belief, according to most everyone, requires concepts. Perhaps it will be argued that some creatures might possess concepts in the sense required for belief without mastering a public language with which to articulate those concepts. But now keep in mind, second, that belief is a rational state, in the sense that the believer is in a position to 'give application' to the question asking for her reasons: either by giving them or by saying she has none. Either way, part of being a believer is understanding the request for a reason, and this cannot be done by a non-linguistic creature. This is not to deny, as I shall emphasize below, that non-rational animals have thoughts about the world. It's just to deny that these thoughts take the same form as ours do.

A final worry about (A): Some argue for the view that to believe is to possess a cluster of cognitive, phenomenal and behavioural dispositions (see Schwitzgebel 2010). If all of the others are present, it might be asked, does it really matter so much if the ability to honestly assert is missing? Could this really make the difference between belief and non-belief? Yes. Given that belief-that-*P*-ish behaviour has sources beyond belief (e.g., visceral fears), such behaviour taken by itself does not constitute evidence in favour of the presence of belief as opposed to one of those other sources. Furthermore, without the connection to what the subject is able to assert, we have decisive evidence that the relevant behaviour is the result

of something *other* than the rational state to which my inquiry is (for the reasons given earlier) confined.

It is worth emphasizing that my point here is *not* that believing is *identical* to the ability to honestly assert. (A) is not offered as a theory of belief, but rather as a truth about belief for which a theory must ultimately account.<sup>6</sup>

(B) S is able to honestly assert that p only if S is able to avow the belief that p.

To assert is to perform a certain act. It is intentionally to put forward the asserted proposition as true (cf. Heck 2006), which is to say that the person who asserts understands that an act of assertion can be criticized if the asserter doesn't in fact believe that p, understands that, in asserting that p, she represents herself as believing that p (cf. Black 1952 and Davidson 2001b, essay 18),<sup>7</sup> as opening herself up to follow-up questions like 'why do you believe that p?' (cf., Boyle 2011). Understanding these things is part of what it is to assert *intentionally* because in the absence of such knowledge, one would not know what one was doing in asserting, one would not be acting under the description 'asserting'.<sup>8</sup>

An assertion is revealed to be less-than-fully honest by a reply such as: 'I'm not at all sure that I *do* believe that p'. If one has any grounds for doubt about whether one believes that p, the assertion is at least somewhat disingenuous. It's not intelligible that someone might honestly assert that the subway is safe, but yet be in a position to be apprised of the fact that she believes that the subway is safe by an appeal to behavioural evidence. Were someone to utter the words 'the subway is safe' and her audience were to respond by saying 'so you actually believe the subway is safe?', and she were to reply 'oh, do I? I hadn't realized', we could no longer consider her initial utterance an assertion at all. It's not as if we would simply continue with something like 'yes, you just asserted that p and assertion that p is an expression of belief, so, assuming you were honest, you *must* believe that p'. Thus it would be

<sup>6</sup> Another objection to (A) comes from an advocate of a knowledge-norm for assertion: one should assert only what one knows. One might argue on the basis of such a principle that if, e.g., one knows that one doesn't *know* that the butler did it, then one cannot honestly assert that he did – even if one believes that he did it. But even if we accept the knowledge-norm, it does not follow that someone who asserts without knowledge is being dishonest. At worst, they are being deliberately misleading. But not every form of deceptiveness is dishonesty.

<sup>7</sup> That an asserter of p represents herself as believing that p is also implied by the view that such an asserter represents herself as knowing that p (see Unger 1975, ch. 6; Slote 1979; DeRose 2002).

<sup>8</sup> See Anscombe (2000) and Davidson (2001a, essay 3) for discussion of the essential tie between doing X and acting under the description 'X'.

bizarre to attempt to *inform* an asserter that she believes the asserted proposition.<sup>9</sup>

Furthermore, it's not just that an asserter of *p* cannot learn that she believes that *p* from someone who witnessed the assertion. Rather, we cannot intelligibly view her as honestly asserting that *p* while only crediting her with an evidence-based awareness of the fact that she believes that *p*. If the supposed asserter were to reply to 'so you actually believe the subway is safe?' with 'yes, I noticed that I believed it right away' or 'yes, it's impossible to miss that fact about myself' we would, again, lose our footing with her original remark. The attitude towards *p* expressed by the assertion is inextricably bound up with the speaker's not having to notice and so not being in a position to miss or not miss the fact that she holds the attitude. The distinctive authority that characterizes avowals of belief is already present in assertion. To accept that the speaker has made an honest assertion is to treat her as being in a position to avow the belief, i.e., to state authoritatively, yet not on the basis of observation or evidence, that she believes the relevant proposition. And so if it is understood that a speaker is able to honestly assert that *p*, a question regarding whether she can avow the belief that *p* is only reasonably interpretable as question about whether there will be an opportunity to do so or whether the ability is masked. By our test, then, the honest-asserting-ability is sufficient for the avowing-ability.

To put this result together with that of the previous section: Believing consists in part in the ability to avow a belief, i.e., to say authoritatively, yet not on the basis of observation or evidence, that one holds the belief. Beliefs are expressible, not just via assertions, but also via avowals of belief.

It might be objected that although the assertion-capable creatures with which we are familiar are able to avow their beliefs, it may yet be merely a contingent truth that asserters are avowers. To defend (B), we would need an argument that necessarily, asserters are avowers.

What rules out an asserting non-avower is implicit in the discussion above: without an ability to ascribe beliefs to itself (authoritatively, but not on the basis of observation or evidence) it could not know what it was doing when it 'asserted' that *p*, i.e., it could not be asserting intentionally. If one were to respond to the 'assertion' of a supposed asserting non-avower by asking 'so, you really believe that *p*?', it would not understand. This lack of comprehension is inconsistent with the hypothesis that it had been intentionally asserting. Furthermore, the notion of

<sup>9</sup> An anonymous referee for *dialectica* suggests a counterexample to my claim that honest assertion expresses belief. Someone in the audience at a horror movie might say 'he's coming to get you!' despite not truly believing anyone was in fact in danger. How precisely we should understand the way in which we are drawn into the world of a fiction while at the same time recognizing that world's non-actuality is a fascinating issue. A full discussion is not possible here. But I would argue that such statements are *assertions* only in the same qualified sense as what they express are *beliefs*.

creatures capable of assertion, but incapable of ever *intentionally* asserting, is nonsensical. Indeed, it is not clear that asserting is something that someone can ever do unintentionally (although one might unintentionally assert that *p*, as opposed to *q*).<sup>10</sup>

It follows from this argument that asserters must possess the concept of belief. Given (A), this means that *belief requires the concept of belief*. This can seem strange. After all, other mental states do not require for their possession the concepts of those very mental states: pain, for example, does not require the concept of pain. Even if belief and assertion require the mastery of a public language, why must this language contain the concept of belief in particular?

It must be because to be intentionally asserting, one must be acting under the description ‘asserting’. And one cannot act under the description ‘asserting’ unless one possesses the concept of assertion. The concept of assertion, in turn, is parasitic on the concept of belief. To assert is, in non-defective cases, to say what one believes. Furthermore, the conceptual connection between assertion and belief must itself be known by the asserter.<sup>11</sup> It’s not just that in asserting, one is doing something that turns out (upon philosophical investigation) to be saying what one believes, in the way that in giving someone a gold ring, one is doing something that turns out (upon scientific investigation) to be giving someone a ring made of an element with atomic number 79. Someone cannot know what she is doing in asserting yet fail to know that she is thereby purporting to say what she believes. This is in part why it would be absurd to assert that *p* and at the same time deny that one believes that *p*, i.e., why Moore-Paradoxical statements are absurd, as opposed to simply revealing a surprising ignorance of the connection between assertion and belief. (It is not credible that someone might, upon uttering a Moorean statement, assuage audience perplexity by saying: ‘Oh, I had no idea that, in asserting that *p*, I was putting myself forward as *believing* that *p*!’) In acting under the description ‘asserting’, one is also acting under the description ‘saying (or purporting to say) what one believes’. Thus, an asserter – and hence a believer, given the argument in favour of (A) – must possess the concept of belief.

The conjunction of (A) and (B) entails that *S*’s believing that *p* by itself suffices for the ability to avow it. However, Freudians might take themselves to have a class of counterexample: deeply repressed beliefs. Let’s reserve the term

<sup>10</sup> It might be objected that we do occasionally make assertions unintentionally. A man might blurt out ‘you lie!’ in a moment of outrage. But such actions, though spontaneous, bear the key marks of being intentional. The speaker gives application to the question that asks for his reasons: ‘why did you say that?’. If interrupted mid-sentence, he does not need to guess (as an onlooker might) how the sentence would have finished. For an action to be intentional, it does not need to be the product of prior planning.

<sup>11</sup> Thanks to an anonymous referee for *dialectica* for pressing me to be clearer on this point.

‘unconscious’ for those. My thesis – that believing entails the ability to avow – would seem to be refuted by unconscious beliefs.

It is undoubtedly natural to think that the difference between an unconscious and a conscious mental state is that one is able to avow the latter but not the former.<sup>12</sup> Still, there is an intuitive sense in which an unconscious mental state is avowable in a way that a subpersonal mental state is not. Therapy can never bring it about that a subject avows her subpersonal states, whose contents the subject need not even be capable of grasping. But even Freudians hold that unconscious mental states could come to be avowed – that’s the goal of therapy. Importantly, ‘able to avow with the help of a therapist’ does not imply ‘unable to avow’. Recall the example of the neurosurgeon above: there is nothing inconsistent about the idea of an ability that one can only exercise with help. Thus, the fact that help from the therapist is needed does not show that, in the absence of such help, the subject is deprived of the ability. One can thus understand unconscious beliefs, beliefs about which the subject is in denial and ordinary beliefs as lying on a continuum of avowability: able to avow with the help of a therapist, able to avow upon the overcoming of denial and simply able to avow. All are avowable. Thus, an unconscious belief is not a true counterexample.

But what of very small children, who are just learning to speak? It might seem, on the one hand, that they are able to assert, and on the other, that they lack the conceptual wherewithal to ascribe beliefs to themselves. I will discuss this objection in section 3 (objection 4) later.

Summarizing the argument for (B): To assert is to put forward a proposition as true. Because this ‘putting forward’ is intentional, it is a putting forward *as* what one believes. Hence, to assert is already to be able to avow a belief.

(C) S is able to avow the belief that p only if S knows that she believes that p.

I think that there’s something right about John Hyman’s account of knowledge, according to which knowledge that p is the ability to do something for the reason that p (Hyman 1999). I’ll rely on a weaker claim, which I have defended elsewhere: knowledge that p is a necessary precondition of acting because p, in the rational sense of ‘because’.<sup>13</sup> The basic idea, briefly, is that in order, say, to walk to the restaurant because *it’s open for lunch*, an agent must know that it’s open for lunch. She cannot do this if it’s not open for lunch or if she believes on the basis of false evidence that it’s open for lunch, although she might do so in such cases because *she believes it’s open for lunch*. To say that she is walking to the restaurant because *it’s open for lunch* is to describe the action as resulting in a certain way from the fact’s being in view.

<sup>12</sup> This is how Finkelstein (2003, ch. 5) describes the difference.

<sup>13</sup> I defend this claim in Marcus (2012, ch. 1).

Now suppose that S holds the belief that p and so, as I have argued, is able to avow it. Suppose, further, that S hears from someone she trusts that those who believe that p tend to have good luck with the online dating service, p-date. There may be barriers of various kinds to S signing up for an account on p-date, but they do not include an inability to act for the reason that she believes that p. That she believes that p can be one of her reasons for signing up for p-date. There may, of course, be obstacles to exercising this ability, either internal or external; and an occasion for exercising it may fail to materialize. But to suppose that she has no such ability cannot be reconciled with the stipulation that she is able to say, authoritatively but not on the basis of observation or evidence, that she believes that p. One cannot both acknowledge the avowal as authoritative in the relevant sense and yet deny the ability to act in light of the avowed psychological fact. Of course, if one thinks of a statement of the form ‘I believe that p’ as a report based on less-than-decisive evidence, then matters are different. But this is exactly not to think of the utterance as an avowal. Speculation on the basis of weak behavioural evidence that one believes that p does not put one in a position to act in light of the fact that one believes that p. But to treat someone as in a position to avow the belief is thereby to credit him with an ability to act on the basis of the psychological fact.

I should emphasize that the point is not that having the ability to avow a belief *puts one in a position to know* that one believes. Insofar as knowledge that p is entailed by having an ability to act in light of the fact, the ability to avow a belief is nothing short of self-knowledge itself.<sup>14</sup>

To summarize: Believing entails that the subject is able to honestly assert, the latter entails that the subject is able to avow the asserted belief, and such an ability entails knowledge that one holds the belief. Therefore, believing that p is inseparable from knowledge that one believes. This thesis is stronger than the more familiar idea that beliefs are self-intimating, understood as the thesis that a second-order belief is a typical causal consequence of a metaphysically independent first-order belief (see Mellor 1977–1978 and Armstrong 1968, ch. 6, section 10). If one is thinking of a state that may or may not involve awareness that one is in that state, one is not thinking of belief. The metaphysical profile of belief has a self-conscious, epistemic dimension.<sup>15</sup>

<sup>14</sup> Moran goes even further, describing “the ability to avow one’s belief as the fundamental form of self-knowledge” (Moran 2001, 150). See also Boyle (2009).

<sup>15</sup> Shoemaker argues for a conclusion in this neighbourhood: “where the subject has the concept of belief and of herself, the first-order belief’s being available *constitutes* her having the at least tacit belief that she has the first order belief” (Shoemaker 1996, 92–93). And Kriegel makes a similar claim about *conscious* beliefs: “conscious beliefs always include an awareness of themselves” (Kriegel 2004, 108). See also Hintikka (1962). It is more common to argue that having a second-order belief constitutes the possession of a first-order belief, see, e.g., Heal (1994, 22): “when I come to think that I believe that p then I do, in virtue of that very thought, believe that p”. See also Kobes (1995).

But the picture of doxastic knowledge that emerges from this argument is remarkable not just in virtue of the necessity of such knowledge, but also on account of its special character. Knowing what one is doing when one asserts, I argued above, requires a non-evidential yet authoritative grasp of one's holding the underlying belief. Since there is no question of *how* (i.e., by what evidence) someone who avows a belief knows she holds it, knowledge that one believes cannot be downgraded to mere belief that one believes because of bad evidence. This helps to explain why a successful challenge to a would-be avowal of a belief registers as the revelation of *dishonesty* rather than *error*.

Some have argued that the inapplicability of epistemic challenge shows that expressive self-ascriptions do not manifest self-*knowledge* at all.<sup>16</sup> If the foregoing arguments are correct, however, it shows only that there are classes of pronouncements that we treat as knowledge-manifesting despite there being no question of their epistemic grounding.

But, it will be objected, if knowing that you believe that *p* involves believing that you believe it, my central thesis seems to imply that any belief brings with it an infinity of ever higher-order beliefs. Part of the idea, however, of the inseparability that I'm proposing here is that the capacity for belief and the capacity for awareness of what one believes are not distinct. The relevant capacity is self-conscious, in the sense that its exercise is a state such that part of what it is to be in the state is to be able to say, not on the basis of observation or evidence but simply in virtue of one's being in it, *that* one is in it. As McDowell says of rational perceptual capacities, we should say that the power to believe "is a single capacity, self-consciously possessed and exercised" (McDowell 2011, 41). This suits the idea that knowing what I believe is not a matter of having checked up on my psychological state. Such checking-up would engage a distinct capacity – the capacity for telling whether one holds a certain belief or not. Ordinary awareness that one believes that *p* and one's believing that *p* itself are a single manifestation of the underlying capacity.

How does this disrupt the burgeoning of ever higher-order beliefs? Consider the third-order belief. Suppose someone were to ask whether I believe that I know that I believe that *p*. What is in question here is not a sort of belief to which my account commits me one way or the other. For it is impossible not to hear this as the question of whether I am in a state about which it would make sense to ask for my evidence – evidence for believing that I know that I believe. And of course, my thesis does not entail the generation of third-order beliefs of this evidential kind. But it's not clear what else there is to make of the idea of a belief that

<sup>16</sup> See Bar-On (2004). Finkelstein warns against assimilating what an avower manifests to ordinary self-knowledge, but is more circumspect on the question of whether it is *really* knowledge (see Finkelstein 2003, 6.4).

I have the relevant sort of doxastic knowledge. There is simply no difference between knowing (in the relevant sense) that I believe that *p* and knowing (in the relevant sense) that I know (in the relevant sense) that I believe that *p*; we have added words without adding meaning (cf. Rödl 2007, 145). All there is is self-conscious belief.

A final puzzle about self-deception: it can seem as if William is in a position to learn of his belief from evidence. How is this possible if he already knows what he believes? How could people ever need evidence to accept what they already know without evidence? And yet it seems that they do. William's mother might ask him, in a quiet moment, to consider the significance of his explosive reaction to the suggestion that the artworld is still really only interested in Alfred, to consider whether this reaction is characteristic of someone who holds the challenged belief or rather of someone who knows the suggestion to be correct. William might feel compelled to acknowledge that he does, after all, agree with his sceptical friends. But then, prior to his mother's intervention, it seems that he must not have known what he believed. Self-deception thus seems to drive a wedge between belief and knowledge of belief.

One way out of this difficulty would be simply to deny that the self-deceived believe the propositions that they would prefer to be false. To explain self-deception would then be in part to explain how it is that such a preference can lead to the belief that the relevant proposition *is* false – and to explain it without imputing to the subject the opposite belief (see, e.g., Mele 1987, ch. 9). This would be to think of William not as believing (let alone knowing) that the artworld remained as uninterested as ever in his work, but instead, perhaps through selective attention and/or motivated reasoning, as arriving at the opposite conclusion. There would then be no belief for William to find out he possessed via evidence, and hence no wedge between belief and knowledge of belief.

However, this strikes me as the wrong way to describe the example (although it is surely the right way to describe others), and so I do not want to rest my argument upon it. Were William to assert that the artworld had finally come to appreciate the value of his work, I would think him dishonest. The question, then, is whether there remains a credible path forward for the view that to believe is to know that you believe.

If one thinks (as I do) that William holds the belief he prefers to be false, then to explain his self-deception is in part to explain how it is that he manages to shield himself from that belief while retaining it. If such shielding amounts to causing himself to be ignorant of the fact that he holds the belief, then this would indeed introduce the separation of what I claim is inseparable. But there is another interpretation. The alternative is that William shields himself not just from the belief he prefers to be false, but also from his own knowledge that he holds the belief. In that case, the shielding could be understood in terms of his directing the stream

of his consciousness away from both the belief and the inseparable knowledge of that belief. Belief is banished from the stream of consciousness, and so is doxastic self-knowledge. Thus, belief and knowledge of belief are not separated after all.<sup>17</sup>

How should we decide which of these two interpretations is correct? I submit that it is perfectly appropriate at this juncture to bring to bear a theory of belief. It is, after all, a very difficult sort of case, one that we naturally find ourselves struggling to describe in a stable way. There is no clear common-sense verdict to which one might appeal in order to settle the matter. I have given an argument for a thesis that entails the second reading, and will give, in the next section, a theory of belief that helps to make sense of this thesis. If the first reading were the only one, William's case would be a clear counterexample. But the second reading is also plausible.

I think this modest dialectical point is sufficient for my purposes here. But there are also compelling independent grounds for preferring the second reading. Consider that the psychological machinery of self-deception is engaged not simply by the offending belief, but by the subject's awareness that she holds the belief. After all, a burglar of whom one isn't aware isn't any more disturbing than the absence of burglars. Similarly, a belief of which one is not aware shouldn't be any more disturbing than a belief that one doesn't hold. But the self-deceived person *is* disturbed by the offending belief. He remains conflicted, even after the deception has been accomplished. He is touchy on the topic of the belief, and apt to fly into a rage when pressed. This sort of reaction is evidence that he holds the belief; but it is equally evidence that he knows that he holds it. I submit that while this is too quick to be decisive, it does suggest a plausible direction for an independent argument to take.

Up until now, I have argued that knowing you believe *p* is part of believing that *p*. In section 3, I will defend a theory of belief that explains this fact.

### 3. *The source of doxastic self-knowledge*

What is it about belief that makes it self-conscious? To believe is to think of a proposition as having a certain feature. What feature? An obvious answer is *truth*.

<sup>17</sup> It might be worried that the subject would then fail to know that she knows that she holds the belief, and so doxastic self-knowledge itself would fail to be self-conscious. However, first, the point of the alternative interpretation is precisely to explain the relevant 'shielding oneself from' in *non-epistemic terms*. Just as this form of self-deception does not lead to ignorance of the belief, it does not lead to ignorance of knowledge of belief. Second, as I argued above, there is in fact no distinction between knowing (in the relevant sense) that I believe that *p* and knowing (in the relevant sense) that I know (in the relevant sense) that I believe that *p*. All there is is self-conscious belief. So there is no separate third-order knowledge of which self-deception could deprive us.

This answer will converge, I think, on the superficially different one I give here. But I will not explore that convergence in this paper. I start from the idea that to believe a proposition is to see it as having some feature in virtue of which it is correct to believe it. Certainly, truth is a feature of a proposition in virtue of which it is correct to believe it. Let's call this feature – which may or may not be the same as truth – *to-be-believed-ness*. To believe a proposition is for the believer to view it as to be believed, to view it as what one ought to believe (cf. Marcus 2012, ch. 1; Boyle 2011, 5.3; Raz 2012, ch. 5).

I don't mean that belief is a state of thinking that one ought to be in that very state. Rather the idea is that, in believing that *p*, one views *p* as meeting a standard that requires cognitive conformity. Belief is the conforming of one's mind in recognition of this doxastic obligation. Recognizing *p* as having this feature is not, according to the view I develop here, an additional belief – the belief that one ought to believe that *p* – but rather simply an attitude towards the relevant proposition. (So this is not the regress-inducing proposal that believing *p* is believing *q*, where *q* is the proposition that one ought to believe that *p*.) One might analogously identify doubting that *p* as having a distinctive attitude towards *p*, as opposed to identifying it with a belief about *p*, viz., that it is doubtful. In thinking of *p* as *requiring* belief, one is recognizing the possibility that such a requirement could fail to be met. But to recognize the requirement is to believe the proposition. In this section, I will elaborate on and defend the idea that what makes an attitude towards a proposition *belief* is that one views it as what ought to be believed.

Viewing a proposition as having a feature in virtue of which one ought to believe it might sound like having *reasons* for believing the proposition. And so it might seem as if I advocate that beliefs are, as such, held on the basis of reasons. But I am not. 'To-be-believed-ness' is what good reasons show a proposition to possess. To regard a proposition as to be believed is thus something that can be done on the basis of reasons. Further, when we consider a proposition to be sufficiently well-supported by reasons, we *should* regard it as to be believed. But regarding a proposition as meriting belief in this sense is logically independent of actually believing it on the basis of reasons. One might so regard it without recognizing any such reasons or recognize such reasons and fail to so regard it. One might even view a proposition as to be believed for which one knows one lacks sufficient evidence, against which one has very good evidence and about which one entertains serious doubts. One thus regards a proposition as meriting belief despite recognizing considerations that appear to show that it doesn't merit belief. These kinds of cases involve irrationality, and so are neither typical nor ideal, but they are nonetheless possible – on my view as much as on our pre-theoretical view.

We shall see below that there is much to be said in favour of this view. For the moment, consider the following two *prima facie* reasons for thinking that belief has an ought-ish character of some sort.

First, when asked to explain *why I do believe p* – at least when this question is understood as asking for my reason – I comply by saying *why one ought to believe p*. Why do I think the butler did it? Because everyone else has an alibi. In so answering, I cast my own belief as the upshot of my recognition of the evidence in favour of thinking of *p* as what to believe. A neat explanation of this fact is that belief is itself just a stance on the question of whether *p* is to be believed.

Second, note what we actually say in agreement or disagreement with an avowal of belief. If I say ‘I believe that *p*’, you object not by saying ‘you don’t believe it’ or by providing evidence that I don’t believe it, but rather by saying something to the effect of ‘you shouldn’t believe it’ or by providing evidence that *p* isn’t true. A straightforward explanation of this fact is that we take believing that *p* to be a stance on the question of whether *p* is to be believed.

One might object to the first datum just adduced as follows: Since we sometimes believe for no reason at all and so cannot say why we believe, there is no reason to think that belief in general has an ought-ish character. But that is wrong: Even when our reply to the question ‘why do you believe *p*?’ is ‘no reason’ or ‘I don’t know’, we recognize that these are weak answers to the question, that these replies put us in the position of maximum vulnerability to contrary evidence. And were anything of importance to hang on the question of *p*, our failing to have or to know our reasons would be grounds for criticism.

I contend my thesis about the nature of belief helps to explain how it could be that to believe is to know that you believe. According to this thesis, there is nothing more to believing *p* than viewing *p* as to be believed. Thus the ability just to say that one *does* believe *p* is nothing beyond the ability to express one’s view of what one *ought* to believe. My explanatory strategy here depends on a certain conception of where the central puzzle about knowledge of one’s own beliefs lies. I take it that it is connected to the absurdity of Moore’s Paradox. It seems mysterious why making a statement about, say, the subway, e.g., ‘The subway is safe’, renders it absurd for me to deny something about myself, e.g., by saying ‘I don’t believe that the subway is safe’. This mystery would be explained if we had some account of why simply having a belief about the subway puts me in a position to state authoritatively some fact about myself. Such an account would show why the first conjunct of Moore’s Paradox seems to reveal the speaker’s knowledge of something that is then denied by the second conjunct. If that’s right, the fundamental issue is what explains (B) above. According to (B), one cannot be in a position to assert that *p* without being in a position to avow the belief that *p*. While there are interesting questions about why it is that believing *p* puts one in a position to assert it and why authoritatively avowing the belief that *p* entails

knowledge of one's belief – (A) and (C) above – these are not confounding in the way (B) is.

But, according to my proposal, in saying that I *do* believe p, I am saying nothing more than that I *ought to* believe p. There is no mystery of self-knowledge in my ability to say that p, given that I believe that p. Similarly, there is no mystery of self-knowledge in my ability to say that p is what I ought to believe, given that I think I ought to believe it. So, if believing p is no more than thinking I ought to believe it, then there should be no mystery about how I am in a position to avow my belief.<sup>18</sup>

This way of putting the explanation, however, courts the misunderstanding I warned against above, viz., that 'I ought to believe p' is the content of a belief, distinct from the content that p. The idea is rather that my recognition of p as what I ought to believe is what makes my attitude towards p *belief*. The attitude towards p that is constitutive of belief, on my proposal, is the recognition of p as what ought to be believed, and so what I in particular should believe. To believe a proposition is to recognize that I myself am bound to believe it. But there should be no more mystery about my knowing that I view p as what I ought to believe than there is mystery about my knowing that it is *doubt* or *incredulity* with which I view the proposition p, where I doubt or am incredulous that p. Similarly, just as my knowledge that it is doubt with which I view a certain proposition does not depend upon evidence, my knowledge that I view p as *to be believed* also does not depend upon evidence. My proposal thus reveals what belief is such that, as shown in section 2, part of what it is to believe is to know non-evidentially that one believes.

I've been describing belief as an *attitude* that a person has towards a proposition. Some analyse attitudes dispositionally. I would reject any such analysis. A crucial defect of dispositional accounts of belief is that they distort the explanatory connection between believing and the behaviour to which believing characteristically gives rise. If, for example, a disposition to assert that p is identified as partly constitutive of believing that p, then the explanatory connection between someone's believing that p and their asserting that p is minimal, comparable to the explanatory connection between a glass's being fragile and its breaking after being dropped. In reply to the question of what it is about a glass's being fragile that explains its breaking after being dropped, one could do no better than say that to be fragile just is to break in such circumstances. Similarly, in reply to the question of

<sup>18</sup> Although I think that a similar account explains how we know, not on the basis of evidence, what we desire and what we are doing (see Marcus 2012, ch. 2), no such account will work for non-evidential knowledge of our own sensations. This is as it should be. Our knowledge of our active, rational attitudes is of a different sort than our knowledge of our passive, non-rational attitudes (see Boyle 2009).

what it is about S's believing that p that explains S's asserting that p, one could do no better than say that to believe that p just is to assert that p under the right conditions. But there ought to be something more to say here. For, intuitively, S is disposed to behave belief-that-P-ishly (and so to assert that p) *because S believes that p*. Belief, in other words, is not identical to but rather explains the relevant dispositions. A theory of belief ought to shed some light on what belief is such that it gives rise the various behaviours, thoughts and feelings that a dispositional theorist correctly views as characteristic of believers. And this is precisely what my theory aims to do. Belief, on my view, is not itself a disposition but is rather the *standing attitude* towards a proposition that is the source or categorical basis of the relevant dispositions.

### 3.1. *Objections*

(1) Insofar as we think of to-be-believed-ness simply as the property of meeting a standard for belief, one might object that no matter what the standard is, believing couldn't simply be the same as thinking of a proposition as to be believed. After all, one might recognize that a proposition met a standard – any standard – but still fail to believe it. This would be something like a doxastic version of weakness of the will.<sup>19</sup>

For example, there might be a certain proposition that one simply couldn't take seriously – say, that one's spouse is an undercover Russian spy. The CIA has, suppose, requested a meeting on the subject and, for prudential reasons, one has decided to take it. The evidence is overwhelming. At the end of the session, one still does not believe that she is a spy, but one is in a state expressible by saying such things as 'I ought to believe that my wife is a spy, but I still can't quite bring myself to believe it'. Such a case appears to show that viewing a proposition as meeting the standard of correctness associated with belief is one thing, actually believing it is another.

I begin my reply by considering an example of deliberation. Suppose a woman is playing poker, trying to figure out whether her opponent is bluffing. She considers his behaviour on prior rounds of betting, which she finds inconclusive. Hoping to get a read on him, she asks him whether he'll show her his hand if she folds. He enthusiastically says he will, thereby revealing that he wants her to fold, a sign that his hand is weak. This evidence settles the matter for her – he's bluffing. Armed by the Russian spy example with the distinction between believing and the judgement that one ought to believe, we can, it seems, break down

<sup>19</sup> Some deny the possibility of epistemic akrasia. See, e.g., Pettit and Smith (1996). If they are right, of course, then the objection doesn't get off the ground.

the transition between deliberation and belief into two steps. There is, first, a step in which she accepts that certain facts bear decisively on the question at hand, the recognition of a binding doxastic obligation: she ought to believe he's bluffing. Second, there is a step in which she actually comes to believe that he's bluffing.

On the face of it, however, this two-step picture is totally artificial. If we really took this to be the relation between deliberation and belief, it would seem that a rational explanation of why a person believes *p* explains directly only why she judges that she ought to believe *p*. An explanation of why she actually believes it would also have to mention that the ought-judgement led to the belief. You ask why she believes he's bluffing and she says 'because he said he'd show me his hand'. You reply 'that's why you think you *ought* to believe he's bluffing, but why do you actually believe it?'. She would be utterly confused. Her original answer seems to leave nothing of this sort to explain. On the two-step view, there would be questions about how precisely the ought-judgement gives rise to the belief, whether it is an automatic process or whether more activity from the thinker is required. If it is automatic, one might ask how long it takes and whether the process can be sped up or slowed down, whether it might still be accomplished while the subject was asleep or drunk. But these are very strange questions. If more activity from the thinker is required, then we must ask what sort of activity this might be – if not more deliberation, then what? Nothing comes to mind. I shall return to this problem with the two-step view in discussing the next objection.

The two-step picture fits the case of the Russian spy. In this case, the husband judges that he ought to believe that his wife is a spy, but can't quite bring himself to do so. If he does later come around to believing it, it may not be in virtue of having acquired any new evidence, but rather only in virtue of taking more seriously the evidence that he already has. In this case, the ought-judgment and the belief are separate matters. The question is whether this shows that there must be some such separation even when there is no other reason for thinking so. I submit that it shows no such thing.

Following what seems like the right thing to say about our two examples, we can make a distinction between two sorts of deliberation: engaged and disengaged. On the one hand, we can take up the question of whether *p* is true in a spirit of doxastic openness; one can give oneself over to the results of one's deliberation come what may. When one deliberates in this engaged manner, one's beliefs are on the line. This is the typical case. If I am considering whether *p* then, upon encountering what I take to be decisive evidence – *q* – in favour of *p*, I thereby believe *p*. There is no extra step necessary, in which I decide whether to adopt the belief supported by the evidence. Nor is there some kind of waiting period, only after which the belief takes root.

On the other hand, one can take up these sorts of questions in a more academic spirit, in a disengaged manner. One might be unable seriously to entertain the idea

that one's spouse is a Russian spy, and yet one might nonetheless examine the evidence. Whether or not one will go along with the results of this inquiry may then not be settled simply by the power of the evidence; one's affection or loyalty may block what one acknowledges are decisive epistemic reasons for belief. This example opens up space in the mind of the deliberator between 'I ought to believe' and 'I do believe' just because he is not engaged in deliberation in the normal way.

Viewing a proposition as 'to be believed' is the characteristic product of theoretical deliberation. More specifically, it is the characteristic product of *engaged* as opposed to *disengaged* deliberation. We might equally well use the same words to express the conclusion of either sort of deliberation: 'I should believe p'. In one case we are looking at the evidence in order to make up our mind what to believe; in the other case we aren't. But only in the former case does the statement express the thinker's representing p as to be believed, in my sense. In the CIA case, the husband is not examining the evidence with the ordinary sort of attitude; his wife's being a Russian spy is not something he is willing to contemplate seriously. Hence, he ends up with the sort of opinion that results from disengaged deliberation, in which a gap opens up between the result of the deliberation and what one in fact believes.

My point is neither that beliefs are always the product of deliberation, nor that engaged deliberation always results in belief, nor that disengaged deliberation never results in belief. My claim is rather that belief is *the characteristic product* of engaged deliberation. There is, on the one hand, thinking that one ought to believe p, where the thinking is of the sort characteristically arrived at on the basis of engaged deliberation; and, on the other hand, there is thinking one ought to believe p, where the thinking is of the sort characteristically arrived on the basis of disengaged deliberation. I am thus using the distinction between the two sorts of deliberation to direct the reader's attention to a distinction between two sorts of normative attitudes that one might take towards a pair of propositions. But neither attitude entails the occurrence of any deliberation.

(2) Believing, then, is a matter of viewing a proposition as possessing a certain normative property, which can be understood as the characteristic product of engaged deliberation – to-be-believed-ness. As I said above, the normative element is not to be understood as a component of the content, but rather as part of the attitude towards p that constitutes belief. So, if propositional attitudes in general can be characterized as ways of viewing a proposition, belief in particular can be further specified as viewing a proposition as what ought to be believed. But one might wonder what this amounts to. In what sense can an attitude towards p be a 'viewing-as-what-ought-to-be-believed', if not by somehow involving a distinct belief, explicit or implicit, with the content 'p ought to be believed'?

I begin my answer to this question where my reply to the previous objection left off: there is no difference between considering whether to believe *p* and considering (in an engaged manner) whether *p* is to be believed. There are not two separate questions, (i) whether to believe the content '*p*' and (ii) whether to believe the content '*p* ought to be believed'. This is what the artificiality of the two-step picture establishes: we do not treat the former question as bearing on a content distinct from the latter question. The question of whether to believe *p* is, necessarily, already the question of whether one ought to believe *p* (in the sense that corresponds to engaged deliberation). The latter merely makes explicit the normative character of the belief-attitude. One does not alter the content under consideration by adding to it 'ought to be believed'. (One might make a similar claim about 'is true'.)

If you think of 'viewing-as-to-be-believed' not as the attitude constitutive of belief, but instead merely as part of the content of a distinct belief, then we are left with a question about how the gap between this normative belief about *p* and the belief that *p* itself is traversed. It's clear why one's rational activity, in the form of weighing the evidence in favour of *p*, can move one to believe that one ought to believe *p*, since the evidence shows precisely what one ought to believe. But how does one go from believing that *p* is to be believed to believing *p* itself, supposing that these are distinct? Not on the basis of more evidence. After all, the evidence in favour of believing *p* is precisely the evidence that one ought to believe *p*. So the transition from the latter to the former (if there must be such a transition) cannot be understood as the direct expression of the subject's rational activity. It would have to be a process of 'sinking in' that one hopes takes place subsequent to one's recognition that *p* deserves to be believed. One could label this process 'rational' or not, but it would amount to viewing the making up of one's mind as outsourced, at the final stage, to unconscious or subpersonal mechanisms that one could at best cheer on as they did or didn't churn out the belief at which one hoped to arrive. Better to reject the gap between viewing *p* as to be believed and believing *p*.

The difficulty for the 'gap' view, I should emphasize, has nothing to do with the immediacy of the alleged transition between believing that *p* ought to be believed and simply believing that *p*. And so it is no use responding that in some cases the latter might be the immediate effect of the former. The point is that, no matter how immediate the transition, belief that *p*, on the 'gap' view, is something external to the agent's doxastic reasoning about *p*, for such reasoning cannot take you beyond believing that *p* ought to be believed.

Against the objection, then, that viewing *p* as-to-be-believed must in the end be understood as a belief with a content distinct from *p*, I have argued that in fact it is simply the belief that *p* itself.

(3) It might be objected that my proposal at best shows that believing something entails knowing that one ought to believe it. How can the gap be bridged between this and knowledge of what one actually does believe? Mustn't it be bridged in part by the subject's knowledge that believing something is recognizing that one ought to believe it? But since this is just a theory, a theory that is not widely held, it is hardly plausible that everyone relies on it whenever they believe anything.

If the theory is correct, however, then we do all already believe it. One point of section 2 is that we already think of non-evidentially knowing what one believes as built-in to believing. Upon hearing someone say 'I believe that p', we do not ask 'how do you know that's what you think?'. And no one (outside philosophy) would know what to make of the question if we did ask it. I have also attempted to make salient to the reader that we answer the question 'why do you believe that p?' in a manner that suggests that we take our believing that p to be dependent on what we think we *ought* to believe. My thesis, then, is that we already operate with an understanding of what it is to hold a belief according to which it is the attitude of taking a proposition to be worthy of belief, and that this explains both why and how we know what we believe. This paper thus aims to bring what we all already know to philosophical self-consciousness.

But it might be argued that even if believers implicitly operate with an understanding of this theory, there is still a gap between thinking of a proposition as to be believed and knowing that one thinks of it that way. How does someone who represents p as to be believed know that she is in fact responsive to the relevant doxastic obligation? How can believing that p enable me to rule out my being incorrect about how I view p? My reply is that it may not, but that this doesn't matter. For my statement 'I believe that p' to be authoritative is for it genuinely to be the expression of my regarding p as to be believed, i.e., my stance on the question of p. If I don't believe it, then, of course, I can neither honestly assert that p nor authoritatively express my belief by saying I believe it. And I might not be able to tell the difference between these cases; my capacity to know what I believe is thus fallible.<sup>20</sup>

The point will be pressed: how could my being in a position to say that p is to be believed put me in the position to say that I represent p as to be believed? To say the former is to say something about p; to say the latter is to say something about myself, which someone else might put by saying 'you believe that p'. However, I am not *saying*, in avowing, that I believe that p in the same sense as I am *saying*, in asserting, that p. Your 'you believe that p' is, like my 'p', an assertion. My 'I believe that p' is not an assertion that I believe that p. Hence 'but it's

<sup>20</sup> Compare McDowell on our self-conscious perceptual capacities in McDowell (2011, 36–44).

false that you believe that p', said in reply to an avowal in an ordinary context, is befuddling and registers as a change of key. My view is an attempt to unpack that metaphor. An avowal 'I believe that p' does not express my representing the proposition that I believe that p as to be believed, rather it expresses my representing the proposition that p as to be believed. It is the verbalized form of the underlying theoretical act, in which I represent the proposition as binding me. Your 'but it's false that you believe that p' treats what I said as if it were an expression of a stance on the truth of 'I believe that p', opposing it with the contrary view. But the proper 'key' of criticism is really not that I've made a false assertion, but that I've spoken dishonestly, since I am not in the avowed state. An authentic avowal is one that expresses the avowed state, and it is authentic avowal that exhibits 'first-person authority'.

(4) According to a fourth objection, my view is refuted by the fact that it cannot be extended to non-rational animals. Many would insist that such animals do believe, despite their lacking the conceptual wherewithal even to grasp propositions, let alone to view them as to be believed. This objection can take several forms; it is worth disentangling them in order to isolate what I take to be the deepest version.

The first version is methodological. By declaring my topic at the start to be *our* beliefs, I might be accused of banishing non-rational animals from the realm of believers by a kind of methodological fiat. But this is wrong. It is not because of how I've elected to use the word 'belief' in this paper that cats, mice and owls cannot give application to the question 'why do you believe that p?'. And it is on the basis of this sort of fact that I conclude that we do not ascribe to non-rational creatures what we ascribe to rational ones using the term 'belief'. Thus, I have not simply defined animal belief out of existence.

The second version is semantic. Someone might say that, according to my view, anyone who uses the term 'belief' to describe a non-human animal is speaking falsely. Hence animal psychologists and zoologists who explain animal behaviour in terms of belief are mistaken. And we should be very reluctant to attribute this sort of widespread error to the experts. However, my thesis is not about the term 'belief', but about what that term normally designates when it is applied to adult humans. The latter states are such that you cannot be in one without knowing it, and this fact is explained by aspects of their nature that also explain why non-rational creatures can't be in them at all. It is perfectly consistent with this thesis that the *term* 'belief' can be correctly used to refer to the states of non-rational creatures.

This brings us to the most worrisome version of the objection. It can seem as if my claim that non-rational animals lack beliefs ultimately commits me to the idea that animals are no more capable of thought than rocks and plants. But plainly, animals learn about their environment on the basis of perception, and act on the basis of knowledge so acquired. Furthermore, they display, to varying degrees, intelligence and ingenuity in deploying their perceptual knowledge. We thus see in animals the same nexus of perception, thought and action that we find in humans, a nexus that is, as I have put it elsewhere, ‘subverted by significant anatomical and genetic overlap and ultimately explained by a shared evolutionary history’ (Marcus 2012, 117).<sup>21</sup> And it can seem as if the only way of accommodating these facts is by holding that the nature of rational and non-rational thought is precisely the same.

Yet we have seen that there are excellent reasons for thinking that we do *not* ascribe to non-humans what we ascribe to humans when we use the term ‘belief’. It is part of the truth-conditions of ‘S believes that p’, where S is a person, that S has the conceptual wherewithal to understand p. Even beliefs about which someone is in denial or beliefs that are repressed in the Freudian sense are such that the believer is capable of grasping the relevant propositions. Belief is thus linked to *understanding*. It could not be otherwise given that believing that p puts one on the hook for answering the question ‘what are your reasons for believing p?’. The requirement that beliefs be eligible to be held on grounds thus goes hand-in-hand with the requirement that the objects of belief be understood by the believer. And it is surely no coincidence that animal thought does not meet either of these requirements and that animals do not possess distinctive first-personal knowledge of their own thoughts. Given these interconnected dissimilarities, it would seem that an account of non-rational thought *should* differ sharply from an account of rational thought. And with our account in hand, many otherwise mysterious aspects of belief have become explicable: (1) why a fully adequate answer to the question of *why I believe p* can be a statement of *why one ought to believe p*, (2) why, in order to believe p, I must have the ability to grasp the proposition that p, (3) why merely holding a belief explains my ability to avow it, and consequently (4) why I cannot intelligibly assert that p and at the same time disavow the belief that p.

The key question is whether one can affirm an account of the sort I advocate without denying the obvious continuities between the minds of non-rational and rational animals. I contend that one can, although there is space here only to sketch the way forward in broad strokes. The core idea is that human thought and non-human thought should be understood as distinct species of a single genus.

<sup>21</sup> I defend at length in Marcus (2013, ch. 3) the view that non-rational animals are not believers.

That both are species of a *single* genus – and so we are not simply equivocating in speaking of *thought* in both cases – is a function of the common role thought plays at the nexus of perception and action in the lives of both kinds of creatures. That there are *distinct* species of thought is a function of the fact that (as we have seen) what one is saying about an adult to whom one ascribes thought is different from what one is saying of a non-rational animal to whom one ascribes thought. On this conception, the acquisition by a species (or an individual) of rationality does not merely add to existing perceptual, cognitive and practical capacities, leaving them as they were (though supplemented by additional capacities). Rather, rationality transforms the original capacities themselves. Thought of any form puts thinkers in cognitive contact with the world, but the form that this contact takes is different in the case of rational and non-rational creatures.<sup>22</sup> The task for a defender of this *transformative* conception of rationality, as Matthew Boyle has called it, is to say, in a detailed and plausible way, what it means to speak in this way of distinct *forms* of thought, as opposed to simply of distinct thought-contents.<sup>23</sup>

Much of what I've just said about animal minds can be said about the minds of very young humans. But special difficulties for my argument might be thought to arise about those who are *becoming* rational. One might suppose that, at a certain stage of development, children possess concepts, but not the concepts that make self-reflection possible, and hence that they assert before being able to self-ascribe. However, this is neither intuitively obvious nor empirically sound.

It should be noted, first, that children start to use mental state words (such as 'think', 'remember' and 'know') to refer to mental states at around age two and a half (Shatz, Wellman and Silber 1983), and that three-year-olds are beginning to acquire the concept of belief – false-belief test evidence notwithstanding (Leslie 1994). So the question comes down to whether children between 18 and 36 months are able to make assertions, in the sense relevant to the satisfaction of the condition specified in (A). An affirmative answer is far from assured by the observation that children at this age are beginning to string words together into sentences. For the question is whether *they* are doing what *we* are doing when we string words together into sentences. *We* are (often) putting forward a sentence as true. And because this 'putting forward' is intentional, it is (as argued above) a putting forward

<sup>22</sup> A full, historical account of human beings would, of course, explain how rational animals evolved from non-rational animals. Such an account is well beyond the scope of this paper. However, it is worth emphasizing that my view does not pose any sort of threat to the possibility of such an account. It does follow from my view that human beings are, in respect of the qualities that constitute their rational nature, qualitatively dissimilar from our ancestors as well as from living species on our phylogenetic tree. But there is nothing odd from the point of view of evolutionary theory about the emergence of qualitative differences.

<sup>23</sup> This understanding of rationality is defended in McDowell (1994). Boyle has gone farther than anyone else in carrying out the project (see Boyle forthcoming).

as what one believes. There is no reason whatsoever to believe that 18- to 36-month-old children do *this*. Any child who possesses the conceptual sophistication necessary to do this would also possess the conceptual sophistication necessary to self-ascribe. Of course this sophistication does not come all at once. As a child slowly masters the network of interrelated concepts that provide the necessary background, her grasp of assertion and avowal are equally incomplete and tentative, and so is her ability to assert and avow (intentionally).

(5) It might be objected that my view has a very odd consequence, namely that in believing that *p*, I think that everyone ought to believe *p*. A problem in this vicinity bedevils accounts of belief according to which one ought to believe that *p* if and only if *p* is true (Bykvist and Hattiangadi 2007). It is worth pointing out that some of these problems do not carry over to my view. Since I am not arguing that all true propositions are to be believed, it cannot be objected to my view that we might as a consequence be obliged to believe true propositions that are impossible to believe truly or that we might be obliged to believe propositions too complex for any of us to believe. The normative character of belief, on my view, is a function not of a norm that links belief to truth, but rather of the fact that believing is itself a normative attitude towards a proposition. A central way believers are governed by norms is written in to the very act of belief itself, in which one's recognition of a norm constitutes responsiveness to it.

But it is nonetheless self-evidently wrong that insofar as I believe that, say, I have only one garbage bag left, I represent this proposition as what everyone should believe. Yet if it is not what everyone should believe, then how exactly can the recognition that *p* is to be believed be at the same time a recognition of a doxastic obligation that binds me, i.e., if not by universal instantiation?

To hold that *everyone* should believe that I have only one garbage bag left would be to hold, absurdly, that a believer would be deficient in some respect for failing to accept this proposition. One might attempt to lessen the absurdity by comparing it to the case of someone who, in a familiar enough sense, ought not to ride a certain elevator because, unbeknownst to him, it is about to get stuck between floors. He might fail to do what he ought, although no one can blame him. Failing to do what one ought, where one is in the dark about facts that would enable one to recognize the obligation, is excusable. One's ignorance can even be completely exculpatory. Perhaps the same can be said for the doxastic obligations I see others as bound by when I believe a proposition to be true.

But the following is still absurd: that in believing that I have only one garbage bag left, I must thereby condemn, excuse or exculpate those who do not share my belief. The revised view no less than the original flies in the face of the fact that in thinking of *p* as what ought to be believed (and so what I should believe), I

recognize an obligation that is relevant to me only insofar as I have taken up the question of *p*. To believe that *p* is to think: with respect to *p*, one should assent, as opposed to deny or sit on the fence. To think of *p* as ‘to be believed’ is, then, to think that everyone who addresses herself to the question of *p* should believe it.<sup>24</sup> But one cannot recognize such an obligation without at the same time knowing that one *has* taken up the question of whether *p*. Hence, to see *p* as to be believed is to see it as to be believed *by me*.

(6) It might seem as if my conception of belief as a standing normative attitude towards a proposition is at odds with an intuitive view of the relation between perception and belief. Believing on the basis of perception is, one might think, simply a matter of taking at face value that things are the way one’s perceptual experience represents them as being. To ‘take at face value’, according to this objection, means that one does not bring to bear any normative concepts, but rather only the concepts corresponding to the content of the relevant belief about the world.

But this paper’s arguments apply equally well to perceptual beliefs, which have the same ought-ish character as beliefs formed by inference or on the basis of testimony. They all take the query: ‘Why do you think so?’. This question might be (somewhat artificially) elaborated upon as follows: ‘Why is it that you regard this proposition as meriting belief?’. And such a question is generally considered to have been satisfactorily answered by ‘I saw it with my own eyes’ or the like.

How precisely to fit belief into a complete picture of human life – exhibiting its connections to perception, non-doxastic cognitive states and action – will have to wait for a more sustained treatment. But here, very briefly, is what I take to be the chief difficulty with conceiving of perceptually based belief as non-normative. So understood, it would be an automatic cognitive effect of undergoing sensory stimulation, something that one found oneself ‘thinking’, just as one finds oneself undergoing perceptual experiences themselves. However this distorts a fundamental contrast between perception and belief, one that Wittgenstein gets at in saying “one can mistrust one’s own senses, but not one’s own belief” (Wittgenstein 1956, 190; cf. Moran 2001, 3.1). It is because perceptual experiences are such that one ‘finds oneself’ undergoing them, that one can trust or mistrust them: one can take them or not take them to be revelatory of the truth. And it is because belief just is the person’s thinking of the proposition as meriting belief that talk of mistrust is out of place. To believe that *p* is to have settled (perhaps only provisionally) the question of *p* for oneself. It is an exercise of the subject’s theoretical agency. But nothing that automatically happens to one can be one’s having settled

<sup>24</sup> This proposal is similar to the one attributed to Wedgwood in Bykvist and Hattiangadi (2007).

any question for oneself (even provisionally). At best it can be a datum that speaks in favour of or against a proposition, a datum that might be cited in justification for believing it or not believing it.

#### 4. Conclusion

Let me return to a remark made above concerning Moore's Paradox. If to believe is to know that one believes, then an assertion that  $p$  expresses not just the belief that  $p$ , but also knowledge that one believes that  $p$ . This fact would go a long way to explaining why we have a difficult time making any sense of statements of the form ' $p$ , but I don't believe that  $p$ '. Interestingly, everything I've said here regarding belief and knowledge of belief goes also for reasons for belief and our knowledge of those reasons. The following conjunction is paradoxical in precisely the same way as Moore's original: 'There are cameras everywhere, so the subway is safe; but I don't believe that the subway is safe because I believe there are cameras everywhere'. There is an argument that stands to this conjunction the way my argument in section 2 stands to the original version of the paradox. One cannot believe that the subway is safe for this reason and yet fail to be in a position to give one's reason in the style of the first conjunct. And one cannot be in a position to so give one's reason and yet be unable to just say explicitly that it is one's reason – i.e., one must be able to avow it. Finally, one cannot be in a position to avow one's reason for belief and yet fail to know what it is. Why should it be that believing for a reason consists partly in knowing that one believes for that reason? Because believing  $p$  for reason  $q$  just is taking the to-be-believed-ness of  $p$  to be a consequence of the to-be-believed-ness of  $q$ .\*

#### REFERENCES

- ALVAREZ, M. 2013, "Agency and Two-Way Powers", *Proceedings of the Aristotelian Society*, **113**, part 1, pp. 101–121.
- ANSCOMBE, G. E. M. 2000, *Intention*, Cambridge, MA: Harvard University Press.
- ARMSTRONG, D. 1968, *A Materialist Theory of the Mind*, London and New York: Routledge.
- BAR-ON, D. 2004, *Speaking My Mind: Expression and Self-Knowledge*, Oxford: Oxford University Press.
- BLACK, M. 1952, "Saying and Disbelieving", *Analysis*, **13**, pp. 25–33.
- BOYLE, M. 2009, "Two Kinds of Self-Knowledge", *Philosophy and Phenomenological Research*, **78**, pp. 133–164.

\*Versions of this paper were presented at Kansas State, Guelph and Auburn University. Thanks to those audiences for their helpful comments. This paper also benefited from discussions with Jason Bridges, David Finkelstein, Anton Ford, Keren Gorodeisky, Matthias Haase, Kelly Jolley, Tom Lockhart, Mark McCullagh, John Schwenkler and Michael Watkins. Gratitude is due especially to Arata Hamawaki for many extended and clarifying conversations. I would also like to thank two anonymous referees and the editors of *dialectica*.

- BOYLE, M. 2011, "'Making up Your Mind' and the Activity of Reason", *Philosopher's Imprint*, **11**, pp. 1–24.
- BOYLE, M. Forthcoming, "Essentially Rational Animals", in: G. Abel and J. Conant, (eds), *Rethinking Epistemology*, Berlin: Walter de Gruyter.
- BYKVIST, K. and HATTIANGADI, A. 2007, "Does Thought Imply Ought?" *Analysis*, **67**, pp. 277–285.
- DAVIDSON, D. 1982, "Rational Animals", *dialectica*, **36**, pp. 317–328.
- DAVIDSON, D. 2001a, *Essays on Actions and Events*, Oxford: Oxford University Press.
- DAVIDSON, D. 2001b, *Inquiries into Truth and Interpretation*, Oxford: Oxford University Press.
- DEROSE, K. 2002, "Assertion, Knowledge and Context", *The Philosophical Review*, **111**, pp. 167–203.
- FINKELSTEIN, D. 2003, *Expression and the Inner*, Cambridge, MA: Harvard University Press.
- GENDLER, T. 2008, "Alief and Belief", *Journal of Philosophy*, **105**, pp. 634–663.
- HEAL, J. 1994, "Moore's Paradox: A Wittgensteinian Approach", *Mind*, **103**, pp. 5–24.
- HECK, R. 2006, "Reason and Language", in: C. Macdonald and G. Macdonald, (eds), *McDowell and His Critics*, Oxford: Blackwell, pp. 22–49.
- HINTIKKA, J. 1962, *Knowledge and Belief*, Ithaca, NY: Cornell University Press.
- HYMAN, J. 1999, "How Knowledge Works", *Philosophical Quarterly*, **50**, pp. 433–451.
- KENNY, A. 1976, *Will, Freedom, and Power*, New York: Barnes and Noble.
- KOBES, B. 1995, "Telic Higher-Order Thoughts and Moore's Paradox", in: J.E. TOMBERLIN, (ed), *Philosophical Perspectives*, Vol. 9, AI, *Connectionism and Philosophical Psychology*, Atascadero, CA: Ridgeview Publishing, pp. 291–312.
- KRIEGL, U. 2004, "Moore's Paradox and the Structure of Conscious Belief", *Erkenntnis*, **61**, pp. 99–121.
- LESLIE, A. 1994, "Pretending and Believing: Issues in the Theory of ToMM", *Cognition*, **50**, pp. 193–200.
- MARCUS, E. 2012, *Rational Causation*, Cambridge, MA: Harvard University Press.
- MCDOWELL, J. 1994, *Mind and World*, Cambridge, MA: Harvard University Press.
- MCDOWELL, J. 2011, *Perception as a Capacity for Knowledge*, Milwaukee, WI: Marquette University Press.
- MELE, A. 1987, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, Oxford: Oxford University Press.
- MELE, A. 2001, *Self-Deception Unmasked*, Princeton: Princeton University Press.
- MELLOR, D. H. 1977–1978, "Conscious Belief", *Proceedings of the Aristotelian Society, New Series*, **78**, pp. 87–101.
- MORAN, R. 2001, *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton: Princeton University Press.
- MORAN, R. 2005, "Problems of Sincerity", *Proceedings of the Aristotelian Society*, **105**, 3, pp. 341–361.
- PETTIT, P. and SMITH, M. 1996, "Freedom in Belief and Desire", *Journal of Philosophy*, **93**, pp. 429–449.
- RAZ, J. 2012, *From Normativity to Responsibility*, Oxford: Oxford University Press.
- RÖDL, S. 2007, *Self-Consciousness*, Cambridge, MA: Harvard University Press.
- SCHWITZGEBEL, E. 2010, "Acting Contrary to our Professed Beliefs or the Gulf Between Occurrent Judgment and Dispositional Belief", *Pacific Philosophical Quarterly*, **91**, pp. 531–553.
- SCHWITZGEBEL, E. 2011, "Knowing Your Own Beliefs", *Canadian Journal of Philosophy*, **35**, supplement, pp. 41–62.
- SHATZ, M., WELLMAN, H. and SILBER, S. 1983, "The Acquisition of Mental Verbs: A Systematic Investigation of the First Reference to Mental State", *Cognition*, **14**, pp. 301–321.
- SHOEMAKER, S. 1996, *The First-Person Perspective and Other Essays*, Cambridge: Cambridge University Press.
- SLOTE, M. 1979, "Assertion and Belief", in: J. Dancy, (ed), *Papers on Language and Logic*, Keele: Keele University Library, pp. 177–190.
- STICH, S. 1978, "Do Animals Have Beliefs?" *Australasian Journal of Philosophy*, **57**, pp. 15–28.
- UNGER, P. 1975, *Ignorance: A Case for Scepticism*, Oxford: Oxford University Press.
- WITTGENSTEIN, L. 1956, *Philosophical Investigations*, Oxford: Blackwell.