

THE REVISABILITY VIEW OF BELIEF

Word Count. 13,157 (inclusive of references and footnotes)

Abstract. I develop and defend the view that subjects necessarily have a certain robust ability to revise their beliefs in response to relevant counter-evidence. Specifically, subjects must be able to revise their beliefs, holding fixed their current psychological mechanisms and skills. If a subject cannot revise a mental state in this way, then the mental state in question is not a belief, though it may be some other kind of cognitive attitude, such as a supposition, an entertained thought, a pretense, or a non-doxastic delusion. The argument for this view moves from the claim that subjects are rationally obligated to revise their beliefs in response to relevant counter-evidence to the conclusion that subjects have a robust psychological capacity to revise their beliefs in response to such evidence. The key to this transition is a certain epistemic version of the principle ‘ought’ implies ‘can.’ Along the way to defending this argument, I defend the view that there are doxastic obligations. I also propose a novel variant of epistemic ‘ought’ implies ‘can’ and argue that this version of the view outperforms traditional versions. Finally, I defend the character-based approach to moral blameworthiness and show how it might be usefully extended to the epistemic domain.

In philosophical discussions of the nature of belief, one of the main controversies is whether subjects are necessarily disposed to revise their beliefs in response to relevant counter-evidence. The claim that subjects are necessarily disposed to revise their beliefs in response to counter-evidence is distinct from the normative claim that subjects *ought* to revise their beliefs in response to counter-evidence. The descriptive claim says that if a subject is not disposed to revise her mental state in response to such evidence, then that mental state is not a belief, though it may be some kind of cognitive attitude, such as an entertained thought, a pretense, or a non-doxastic delusion.¹

The view that subjects are necessarily disposed to revise their beliefs in response to relevant evidence faces the following difficulty: subjects often form beliefs in response to highly inadequate evidence, and they regularly retain beliefs which are contravened by excellent evidence (Mandelbaum 2014, Nickerson 1998). Thus, accepting the view

¹ Those who think subjects are necessarily disposed to revise their beliefs include: Adler (2002), Currie & Ravenscroft (2002), Egan (2009), Gendler (2008), Shah & Velleman (2005), and Velleman (2000). Those who reject this view include: Bayne & Pacherie (2005: 183), Bortolotti (2011: 124), Gertler (2011), Huddleston (2012), Mandelbaum (2014), and Viedge (2016). In addition, the interpretative view of the mind of the kind associated with Davidson (1984) and Dennett (1989) entails that subjects are necessarily disposed to revise their beliefs in some circumstances, though it is not clear exactly which circumstances these are.

that subjects are necessarily disposed to revise their beliefs risks saddling us with the implausible result that beliefs are rare in human psychology.

In light of the of the preceding worry, I think we should reject the view that subjects are necessarily disposed to revise their beliefs in response to relevant evidence. At the same time, there is reason to retain a different connection between belief and rational revision. In this paper, I develop and defend the view that subjects necessarily have a certain robust *ability* to revise their beliefs in response to relevant counter-evidence. Specifically, subjects must be able to revise their beliefs, holding fixed their current psychological mechanisms and skills. If a subject cannot revise a mental state in this way, then the mental state in question is not a belief, though it may be some other kind of cognitive attitude, such as a supposition, an entertained thought, a pretense, or a non-doxastic delusion. I dub the view that beliefs are subject to this constraint *the revisability view of belief* (or *the revisability view*, for short).

Since the revisability view does not require that subjects form beliefs in a rational way, the view is compatible with empirical evidence that subjects frequently form beliefs in response to merely entertained propositions, whether or not those propositions are evidentially supported. The view can also accommodate evidence that in many cases, subjects fail to revise their beliefs in response to relevant counter-evidence. So long as subjects have the relevant ability to revise these mental states, these states can count as beliefs.

The fact that the revisability view can accommodate irrational beliefs shows merely that the view crosses a hurdle which any view of belief must ascend. It does not suggest a positive reason to accept the view. The centerpiece of the paper is such an argument, which I call *the argument from the norm of revision*. This argument moves from the claim that subjects are rationally obligated to revise their beliefs in response to relevant counter-evidence to the conclusion that subjects have a robust psychological capacity to revise their beliefs in response to such evidence. The key to this transition is a certain epistemic version of the principle ‘ought’ implies ‘can.’ Painted in broad strokes, the argument is as follows:

- (1) If some subject believes p and has sufficiently strong, undefeated evidence that $not-p$, then that subject has a *prima facie* obligation of rationality to revise that belief.
- (2) If some subject has a *prima facie* obligation to revise one of her mental states, then that subject has a certain robust psychological ability to revise that mental state.
- (3) If some subject believes p and has sufficiently strong, undefeated evidence that $not-p$, then that subject has a certain robust psychological ability to revise that

belief.

Three points about the argument are worth highlighting at the outset: First, the reason the obligation described in (1) is merely *prima facie* and not all-thing-considered is that obligations of morality or of prudence might in some cases trump this obligation.

Second, the argument's conclusion does not entail that humans in fact possess beliefs. Rather, the argument's conclusion is in principle consistent with this result that humans entirely lack beliefs, for the reason that they lack the ability to revise any of their mental states in the relevant way. However, I will argue that humans can revise a wide range of their mental states in the relevant way and thus, that the revisability view does not leave the class of belief poorly populated.

Finally, the argument's conclusion is consistent with the possibility that subjects might possess beliefs which are never contravened by the evidence. The subject in such a state trivially counts as capable of revising that state, in virtue of never failing to revise that state in response to relevant counter-evidence.

Here is the plan for the paper: I begin by elaborating on the revisability view and drawing out some of its consequences, focusing in particular on how the view can accommodate mental states which are never contravened by evidence (§1). I next develop and defend the main argument for the revisability view (§2). I then consider and reject a natural objection to the view, one derived from the fact that the view excludes from the class of belief certain mental states which are sincerely endorsed, action-guiding, and inferentially promiscuous. (§3). I close with a brief summary (§4).

Before proceeding to the first section, a methodological point is in order. The proposed account is not a theory of how reasoners *in fact* use the concept of belief. Rather, the account attempts to identify what is at the core of the concept of belief. To this end, I subject several platitudes about belief to scrutiny. I find that the platitude that beliefs are governed by certain norms of rationality survives scrutiny. I find that certain other platitudes about belief—such as the platitude that mental states which are inferentially promiscuous are necessarily beliefs—wither under scrutiny.

It should further be noted that while this project amounts to a kind of tidying up of the concept of belief, this tidying up is not a matter of stipulation, nor is it conducted in a self-consciously pragmatic way. Rather, this tidying up is meant to reflect considerations which are broadly accessible. The outcome of this conceptual tidying up is a moderately revisionary approach to belief: while many states we supposed to be beliefs are beliefs, at least some mental states which ordinary usage would count as beliefs are not beliefs.

It is also worth clarifying that this paper has two different aims. The first is to show that there are mental states which are revisable in a certain way and that these

states are of philosophical import. The second is to claim that beliefs are a (possibly proper) subset of this class of mental states. Going forward, I tend to conflate these aims, but they are in principle mutually dissociable.

Finally, it is worth saying something about the import of this paper's thesis for issues outside of epistemology. If this thesis is true, it may have implications for certain recent disputes about mental ontology. These disputes concern the nature of several theoretically vexing mental states, such as clinical delusions, disavowed prejudices, and religious convictions. In debates about such states, the question invariably arises whether the states in question are beliefs or whether they belong to some other kind, such as pretense. Often, those who argue against treating such states as beliefs argue that subjects are not disposed to revise these states in response to evidence, whereas subjects are necessarily disposed to revise beliefs in response to evidence.²

Since I reject the view that subjects are necessarily disposed to revise their beliefs in response to evidence, I side with those who think this condition cannot be used to evaluate whether clinical delusions, disavowed prejudices, or religious convictions are beliefs. At the same time, I think the revisability view has the potential to reconfigure disputes about these states. Instead of asking whether subjects are necessarily disposed to revise such states, we should ask whether subjects have the relevant psychological ability to revise such states. If subjects lack this ability, then these states are not beliefs. If subjects have this ability, it remains a live possibility that these states are beliefs.

1 The Revisability View of Belief

In this section, I elaborate on the revisability view and draw out some of the view's consequences. Here is the full view:

THE REVISABILITY VIEW OF BELIEF

Necessarily, if some subject's mental state that p is a belief and if that subject has sufficiently strong, undefeated evidence that *not-p*, then that subject is able to

² For the view that clinical delusions are not beliefs, see Currie & Jureidini (2001), Currie & Ravenscroft (2002), Frankish (2012), Mullen & Gillett (2014), and Stephens & Graham (2004). For the view that clinical delusions are beliefs, see Bayne & Pacherie (2005) and Bortolotti (2009). For the view that some delusions are at least partly made up of beliefs, see Radden (2010, 2014). For the view that disavowed prejudices are not beliefs, see Levy (2015), Madva (2015), Madva & Brownstein (2016), and Schwitzgebel (2010). For the view that disavowed prejudices are beliefs, see Mandelbaum (2016). For the view that religious convictions are not beliefs, see Van Leeuwen (2014, 2016, 2017). For the view that religious convictions are beliefs, see Boudry & Coyne (2016) and Levy (2017).

revise that mental state, holding fixed that subject's current psychological mechanisms and skills.

Notice that the view says that a subject can necessarily revise her belief, given the psychological mechanisms and skills she already has. Thus, it is not enough for a state to qualify as a belief that its subject could revise it, were that subject to acquire new psychological mechanisms or skills. (I elaborate on psychological mechanisms and skills in §2.2).

For some subject to revise one of her mental states, she must reduce the strength of that mental state, suspend that mental state, or eliminate that mental state altogether.³ While this revision must be something the subject does, it need not be under the subject's voluntary control. In this way, belief revision might be like many other activities which subjects do but which aren't under their voluntary control. These include recognizing their loved ones on sight, falling asleep, and distinguishing between green and blue.⁴

Importantly, the claim that a given subject can revise any of her beliefs does not entail that this subject can *simultaneously* revise all of her beliefs. This restriction is consonant with how we think of other abilities: the fact that Janelle can swim a mile and can also play the *Rhapsody in Blue* clarinet solo does not entail that Janelle can swim a mile *while* playing the *Rhapsody* solo.

It is worth stressing that the revisability view is not a full account of belief. Rather, the view posits only a necessary condition on belief, rendering it consistent with several other proposed necessary conditions on belief. For instance, the revisability view is consistent with the claim that subjects are necessarily disposed to act in accordance with their beliefs (Schwitzgebel 2001, 2002, 2010). The view is also consistent with the claim that beliefs are necessarily *inferentially promiscuous*, or available as a premise across a wide range of their subject's psychological inferences (Glüer & Wikforss 2013b, Mandelbaum 2014, Stich 1978: 507). At the same time, the view is at odds with the suggestion that behavioral dispositions or inferential promiscuity are *sufficient* for belief. (I defend this implication of the view in §3).

One question that arises about the revisability view is how it treats cases in which a subject who has a mental state that p never comes into possession of evidence that *not-p*. There are two ways this case might be spelled out, and the revisability view treats these versions differently. On the first way the case might be spelled out, it is a merely contingent fact that the subject never encounters relevant counter-evidence. If this subject could revise this mental state, were she to possess relevant counter-evidence, then

³ The separate mention of belief suspension is in deference to Friedman's argument that belief suspension is distinct from both belief elimination and shifts in belief strength (2013).

⁴ These examples are from Chuard & Southwood (2009) and McHugh (2012). See footnote 15.

the state can enter into the class of belief. If this subject could not revise this mental state, even if she had relevant counter-evidence, then this state is not a belief.

On the second way the case might be spelled out, it is a matter of necessity that the subject lacks counter-evidence for her view. Perhaps mental states which represent obvious conceptual truths are of this variety, such as the state which represents *nothing is round and square*. Or, perhaps the mental states of omniscient beings are of this sort. The revisability view can permit all such states (if any) into the class of belief, in virtue of the fact that subjects never fail to revise these state in response to counter-evidence. That is, subjects are trivially capable of revising such states.

It has already been suggested that the revisability view can permit irrational states into the class of belief. By way of further supporting this claim, consider how the view treats mental states which are sustained by *confirmation bias*. Confirmation bias occurs when one selectively attends to or inflates the importance of evidence which supports one's pre-existing view (Nickerson 1998). Mental states which are sustained by confirmation bias are irrational and yet, the revisability view can count at least many such states as beliefs. This is because subjects can revise such states. For instance, in one study, it was demonstrated that subjects who held a certain view tended to focus on evidence which confirmed that view, thus exhibiting confirmation bias. Nevertheless, when these subjects were encouraged to consider evidence which might disconfirm their view, they did so, and as a result, they became less confident in their original view (Schwind et al. 2012).

I turn next to the argument for the revisability view.

2 The Argument from The Norm of Revision

The argument in favor of the revisability view is *the argument from the norm of revision*. This argument is meant to extend in full generality to all doxastic states, whether occurrent or dispositional, attended or unattended, unconsidered or reflectively endorsed, conscious or non-conscious, compartmentalized from other states or integrated with other states, heuristically-produced or inferentially-produced. Thus, the argument is meant to extend both to beliefs and to what are sometimes called judgments.

The argument is named after its first premise, the claim that subjects are rationally obligated to revise their beliefs when those beliefs are contravened by the right sort of counter-evidence. Since this claim is normative, it cannot by itself illuminate the descriptive nature of belief. But combining this claim with an epistemic variant of the principle 'ought' implies 'can' yields an argument in favor of the revisability view. Here is this argument:

THE ARGUMENT FROM THE NORM OF REVISION

- (1) Necessarily, if some subject believes p and has sufficiently strong, undefeated evidence that $not-p$, then that subject has a *prima facie* obligation of rationality to revise that belief.
- (2) Necessarily, if some subject has a *prima facie* obligation of rationality to revise some mental state, then that subject can, holding fixed those psychological mechanisms and skills which she already possesses, revise that mental state.
- (3) Necessarily, if some subject believes p and has sufficiently strong, undefeated evidence that $not-p$, then that subject can, holding fixed those psychological mechanisms and skills which she already possesses, revise that belief.

Since the argument extends to arbitrarily chosen subjects and to arbitrarily chosen beliefs, the conclusion is that for any given subject and for any of that subject's beliefs, that subject has a certain robust ability to revise that belief. If this subject lacks this ability, then the state in question is not a belief, though it may be some other cognitive attitude, such as an assumption, a merely entertained thought, or a pretense.

2.1 Does This Style of Argument Prove Too Much?

Each of the argument's premises are controversial, and I will defend each momentarily. Before doing so, it is important to exclude a different objection, one rooted in the suspicion that this *style* of argument cannot succeed, as one might develop similar arguments which have implausible consequences. Consider, for instance, the following argument, which is similar to the argument from the norm of revision, except that it is centered around a norm regulating the formation of belief, instead of the revision of belief:

THE ARGUMENT FROM THE NORM OF FORMATION

- (4) Necessarily, subjects ought to form their beliefs in response to good evidence.
- (5) Necessarily, if some subject ought to form some mental state in response to good evidence, then that subject can form that mental state in response to good evidence.
- (6) Necessarily, subjects can form their beliefs in response to good evidence.

The problem with this argument is that its conclusion can be combined with an empirically plausible view about belief formation to yield the result that beliefs occur only very rarely in human psychology. The view to which I refer is the view—already mentioned in the paper’s introduction—that humans by default believe all or nearly all of the propositions they entertain, whether or not those propositions are supported by the evidence (Mandelbaum 2014). If this view is correct, then many and perhaps most beliefs are not *formed* in response to good evidence, even if those beliefs can be revised in response to good evidence. Thus, pairing the argument’s conclusion with this view of belief formation forces us to the surprising result that beliefs occur only very rarely, if at all, in the minds of humans.

I think we should reject any view which saddles us with the result that beliefs occur only rarely, if at all, in the minds of humans. That is, the platitude that beliefs are common in human psychology should take precedence over other platitudes about belief. For, it is precisely because we assume that beliefs are common in the minds of humans that we have an interest in the concept of belief.

Thus, my diagnosis of the argument from the norm of formation is that it is valid, but its first premise is false. So, the argument is unsound. The norm of formation turns out to be false precisely because it would force us to the view that there are no or very few beliefs in the minds of humans. Put otherwise, the so-called norm of formation is not really a norm at all. It is a pseudo-norm, i.e., a proposed norm which has all the initial appeal of a genuine norm but whose implications render it unacceptable.⁵

Unlike the norm of formation, the norm of revision does not force us to the unacceptable result that beliefs occur only rarely in the minds of humans. For, even if humans very often form views in response inadequate evidence, humans can revise a very wide range of their views when those views are contravened by the evidence. Consider, for instance, the previously mentioned result that subjects moderate their views when they are encouraged to consider evidence which runs counter to their views (Schwind et al. 2012).⁶

⁵ If this is right, then norms of rationality are constrained by empirical facts, including—in at least some cases—empirical facts which are discoverable only via experimental psychology. This means that in at least some cases, settling whether some proposed norm of rationality is genuine requires that we engage with empirical work on human cognition. I welcome this result.

⁶ Further evidence that humans can revise a wide range of their mental states comes from evidence that the mechanism of cognitive dissonance governs a very wide range of mental states and that this mechanism helps subjects rationally revise those states (Harmon-Jones 2000, Harmon-Jones & Harmon-Jones 2007). There is also empirical evidence that subjects can rationally revise heuristically-produced judgments, or what are sometimes termed *System 1 outputs* (*removed for anonymous review*).

Thus, I fully acknowledge the existence of arguments which are unsound and which are structurally similar to the argument from the norm of revision. At the same time, I deny that the failure of these cousin arguments gives us a reason to doubt the argument from the norm of revision. For, what goes wrong with these structurally similar arguments is that they appeal to pseudo-norms and hence, rest on a false premise. In contrast, the premises in the argument from the norm of revision are both true, as I will now argue.

2.2 *The Norm of Revision*

The first premise of the argument is the norm of revision. It says that the subject who believes p and who has sufficiently strong, undefeated evidence that *not- p* ought to revise her belief. Here, what kind of evidence is sufficiently strong will depend on the ultimate correct theory of how subjects ought to update their beliefs in the light of evidence.

The obligation in question is *prima facie* in the sense that it can be trumped by other obligations—whether those obligations are obligations of rationality, morality, prudence, or something else. Though *prima facie* obligations can be trumped, they do not cease to exist when trumped. Rather, such obligations retain their force even when they are not, all things considered, what their subject ought to do.⁷

In favor of the norm of revision is that it explains several ordinary but powerful intuitions. For instance, it seems that if you believe there is fruit on your kitchen table and then, walking into the kitchen, see that the fruit bowl is empty, you should revise your belief. It seems that if you believe your child did not shoplift from a local convenience store and then view surveillance footage showing her doing just that, you should revise your belief. It seems that if you believe that Missoula is the capitol of Montana and then are informed by a native Montanan that Helena is Montana's capital, you should revise your belief.

⁷ One natural question about the norm of revision is whether it is *wide-scope* or *narrow-scope*, i.e., whether the obligation to revise one's beliefs is in some sense in effect even for subjects who lack relevant counter-evidence, or whether the obligation is not in effect for subjects until relevant evidence obtains. It is controversial whether norms of rationality should be construed as wide-scope or narrow-scope and relatedly, what the implications of these different construals are. See, e.g., Kolodny (2005), cf. Broome (2007). I am neutral on the question of whether the norm of rationality is wide-scope or narrow-scope, preferring to let the more general question settle the matter. If it should turn out that norms of rationality are wide-scope, then the norm of revision will also be wide-scope. If it should turn out that norms of rationality are narrow-scope, then the norm of revision will also be narrow-scope.

The claim that you have a rational obligation in such cases is consistent with the claim that morality or prudence might recommend different courses of action. For instance, it might be that morality requires that you believe your child when she says that she did not shoplift, even if the surveillance footage says otherwise. The presence of such an overriding moral obligation would not make it the case that you no longer have a *prima facie* obligation of rationality to revise your belief. In such a case, the moral obligation merely trumps the obligation of rationality such that the obligation of rationality is not, all-things-considered, what you ought to do.

Importantly, the norm of revision extends in full generality to all beliefs, including non-conscious beliefs and beliefs formed on the basis of perception. Suppose you judge, on the basis of a perceptual experience, that two lines in a figure are of different lengths. If you should come to possess good evidence that this figure is illusory in just this respect, you should revise your belief. Or, suppose you come to believe implicitly, as the undetected effect of watching too many commercials, that drinking fruit juice is healthy. If you should come to possess good evidence that fruit juice has deleterious effects on health, you should revise your belief.

Despite the intuitive appeal of the norm of revision, there are several theoretical reasons to reject it. One such reason comes from the view that if beliefs are not under voluntary control, beliefs cannot be governed by obligations. Another reason comes from the view that norms of rationality, if any, must be action-guiding, but that no putative norm of rationality is action-guiding. I will consider and reject each of these reasons for rejecting the norm of revision.

2.2.1 An Objection from The View That Obligations Are Under Voluntary Control

The first objection to the norm of runs as follows: Activities governed by obligations must be under voluntary control. Beliefs are not under voluntary control. So, beliefs are not governed by obligations. Hence, the norm of revision, which posits an obligation concerning belief, is false (Alston 1988).

My strategy here is to remain neutral on the question of whether beliefs are under voluntary control.⁸ Instead, I will criticize the claim that activities governed by obligation must be under voluntary control. I will consider and reject two motivations for endorsing this claim.

2.2.1.1 The View That Blameworthiness Entails Voluntary Control

One reason to think that obligations must be under voluntary control derives from a plausible view of the relation between obligations and blameworthiness. On this

⁸ For a defense of the view that beliefs are not voluntary, see Ryan (2003). For a defense of the view that beliefs are voluntary, see Steup (2008).

view, if some subject fails to carry out an all things considered obligation, then she is blameworthy for that failure. If this is right and if activities which are obligated are sometimes not under voluntary control, then there will be cases in which subjects are blameworthy for activities not under their voluntary control. This result has seemed unacceptable to many (Rusin 2010).

I will argue that we should reject the initially appealing view that subjects are blameworthy only for those activities which are under their voluntary control. In its place, I suggest we opt for a view on which subjects are blameworthy only for those activities which reflect on their *character*, i.e., on their desires, values, or cares.⁹ On this view, there can be obligations of belief even if beliefs are not under voluntary control. The sense in which one can be rationally blameworthy for failing in these obligations is that these failures reflect on one's rational character, i.e., on whether and how much one desires, values, or cares about epistemic goods, such as truth.

The view that subjects can be blameworthy only for those actions under their voluntary control is motivated by intuitive contrasts between certain cases. If you kick someone out of spite, you are blameworthy for doing so. In contrast, if you kick someone merely because you are having a seizure, you are not blameworthy for doing so. One natural explanation of this difference is that the spiteful kick is under voluntary control, whereas the seizure-driven kick is not. This suggests the view that one can be blameworthy for some activity only if that activity is under voluntary control. Relatedly, some activity can be governed by obligations only if that activity is under voluntary control.

Though the voluntarist explanation of blameworthiness can account for the difference in the 'kicking' cases, it is not the only view which can do so. On an alternative view, what explains the difference in these cases is that the spiteful kick reflects poorly on one's character—i.e., on one's desires, values, or cares—whereas the seizure-driven kick does not. The spiteful kick suggests that one does not care enough about others' well-being enough, whereas the seizure-driven kick says nothing whatsoever about what one cares about.

If both the voluntarist and character-based views of blameworthiness can explain the contrast between cases like the 'kicking' cases, we need some further way of adjudicating between the views. Toward this end, I will draw on the work of several theorists

⁹ What I am calling the character-based view of responsibility (here used to explain blameworthiness) is sometimes termed *attributionism*. This view sub-divides into several variants. For instance, the desire-based, value-based, and care-based approaches—which I treat as interchangeable—are in fact distinct and competing approaches.

Main sources of the character-based view include: Arpaly (2004), Frankfurt (1971, 1988), Jaworska (1999, 2007), Sher (2009), Shoemaker (2003, 2013), Smith (2005, 2008, 2012), Sripada (2010), and Watson (1975, 1996). More recent treatments of the character-based view include Brownstein (2016) and Sripada (2015).

who argue that there are obligations of emotion, that flouting these obligations makes one blameworthy, and that these obligations arise whether or not emotions are under voluntary control (Adams 1985; Chuard & Southwood 2009: 614-625; McHugh 2012: 87-89).¹⁰ I will then show that while the voluntarist view of blameworthiness cannot accommodate obligations of emotion, the character-based view can.

Intuitively, you should feel outrage in response to deep injustice, gratitude for undeserved kindness, and remorse when you have needlessly hurt others (Adams 1985; Card 1988; Chuard & Southwood 2009: 614-625; Hieronymi 2008: 362; McHugh 2012: 87-89). If we take these intuitions at face value, there are obligations to feel certain emotions in certain circumstances.

There are multiple strategies one might use to attempt to explain away intuitions that there are obligations of emotion. For instance, one might claim that what is 'off' about the person who (say) fails to feel remorse for harming others is that the person who exhibits such callousness is unlikely to be a trustworthy person. Thus, this lack of remorse is a sign of poor character, but it violates no obligation (McHugh 2012: 87). On another strategy, one might claim that what the relevant cases show is merely that there are sometimes obligations to act as though one has certain emotions. For instance, when one hurts a friend, one should act as though one is remorseful, but there is no obligation to in fact feel remorse (Adams 1985: 4).

As Conor McHugh has argued, the suggestion that the relevant cases can be explained without positing obligations of emotion is hard to square with intuitions about particular cases. For instance, the person whose friend feels no remorse for hurting her might be justifiably more injured by her friend's callousness than by her friend's initial wrongdoing. Moreover, this person might justifiably blame her friend more for her lack of remorse than for the initial wrongdoing. Thus, her friend's lack of remorse *itself* seems to be blameworthy and *itself* to go against what friendship requires (McHugh 2012: 87; see also Adams 1985: 4-6). Likewise, were a wronged friend to discover that her friend's apparent remorse had been feigned, she might justifiably be hurt by her friend's failure to feel genuine remorse. This suggests that merely acting as though one feels remorse does not satisfy the obligation in question. One must in fact feel remorseful.

It would seem then, that we have some reason to posit emotions of obligation, obligations which, when flouted, can result in one's being blameworthy. At the same time, it is at least doubtful whether emotions are under voluntary control. Suppose someone offers you a large financial reward to feel gratitude for a kindness shown to you by a person you have long loathed. Could you collect the money? It seems at least doubtful whether you could. In contrast, suppose someone offers you a large financial reward to

¹⁰ Similarly, Alex King (2014) argues that there are obligations to have certain motives, and that motives are not under voluntary control.

raise your right hand. Could you collect the money? It seems overwhelmingly likely that you could.¹¹ The difference between these cases suggests that it is not clear whether emotions are under voluntary control.

It might be suggested that even if emotions are not under the same sort of control as bodily actions, emotions might be controlled by other, indirect means. For instance, perhaps by focusing your attention on the benefits conferred on you by your enemy's kindness, you might coax yourself into feeling gratitude to her. However, even if this sort of indirect strategy might succeed in some cases, there are presumably other cases in which it is bound to fail. Extenuating circumstances and the psychology of the subject in question both seem to play a role in whether indirect manipulations will have any chance of success.

Here, then, we have two facts: First, there is the fact that there are obligations of emotion. Second, there is the fact that it is *not clear whether* emotions are under voluntary control. I take the combination of these facts to support the following claim: our intuition that there are obligations of emotion is not hostage to the view that emotions are under voluntary control. For if it were, given the lack of clarity about whether emotions are under voluntary control, we would not have this intuition.

My claim is the following counterfactual one: were it to turn out that emotions are not under any kind of voluntary control and were we to understand this fact, this would do nothing to diminish the conviction that you should feel outrage at deep injustice, gratitude for undeserved kindness, and remorse for needlessly hurting others. I take this (counterfactual) fact to be best explained by another (counterfactual) fact: there would still be obligations of emotion, even if emotions were not under voluntary control. Notice that this counterfactual claim is enough to falsify the view that activities governed by obligations are necessarily under voluntary control.¹² On the presumption that failing to satisfy an obligation can make one blameworthy, this result also shows that one can be blameworthy for activities not under one's voluntary control.

It is in accounting for obligations of emotion that the character-based view of blameworthiness outperforms the voluntarist view. For the character-based view can explain why emotions can be governed by obligations—obligations which, when flouted, result in blameworthiness—even if emotions are not under voluntary control. For instance, feeling remorse for hurting others reflects that one cares about others' well-being.

¹¹ This common strategy of testing an activity's voluntariness is due to Alston (1988: 263)

¹² Here my argument departs from both Chuard and Southwood's and McHugh's, who argue that there are obligations of emotion despite the fact that emotions are *in fact* not under voluntary control (McHugh 2012: 87-88; Chuard & Southwood: 2009: 620). My argument relies on the strictly counterfactual claim that if emotions were not under voluntary control, there would still be obligations of emotion.

Feeling outrage in response to a minor slight betrays that one cares too much about one's own respect. In this way, emotions can reflect on the moral character of a person, even if those emotions are not under voluntary control.¹³

Once we have the character-based view of blameworthiness in hand, we are in a position to explain how beliefs might be governed by obligations even if beliefs are not under voluntary control. This is because beliefs can reflect on one's *rational character*, i.e., on whether and how much one values, desires, or cares about epistemic goods, such as truth and knowledge. For instance, it reflects positively on one's rational character if one believes that whether some coin will come up heads does not depend on whether that coin came up heads in the previous toss. It reflects negatively on one's rational character if one believes that astrology is a good predictor of future outcomes. One can be praiseworthy for the first belief insofar as it reflects that one cares about truth or the other epistemic goods. One can be criticized for the second belief insofar as it reveals that one doesn't care enough about truth or the other epistemic goods.¹⁴

I conclude that considerations from blameworthiness do not support the view that there are no obligations of belief.

2.2.1.2 *The View That 'Ought' Implies Voluntarily 'Can'*

I now turn to a different argument that if beliefs are not under voluntary control, then there are no obligations of belief and thus, that the norm of revision is false. This argument runs as follows: first, 'ought' implies 'can,' i.e., if some subject ought to perform some activity, then that subject can perform that activity. Second, 'ought' implies 'can' entails or should be construed as the claim that 'ought' implies *voluntarily* 'can,' i.e., if some subject ought to perform some activity, that then subject can voluntarily perform that activity. Thus, if beliefs are not under voluntary control, they cannot be subject to obligations. Hence, the norm of revision, which posits such an obligation, is false.

¹³ The character-based approach is not the only non-voluntarist approach to obligation. On the reasons-responsiveness approach, activities governed by obligations are responsive to reasons in a certain way, whether or not those activities are under voluntary control (see, e.g., Hieronymi 2008, 2014 and McHugh 2012: 89-93, 2013. McHugh thinks reasons-responsiveness is not a kind of voluntary control, whereas Hieronymi thinks it amounts to a kind of voluntary control). I have said nothing to exclude the reasons-responsiveness approach, and for present purposes, I do not need to. The character-based and reasons-responsive approaches both permit that beliefs might be governed by obligations, even if beliefs are not under voluntary control.

¹⁴ Note that the fact that someone is blameworthy does not entail that it is appropriate for others to censure that person. Indeed, I suspect that the cases in which it is appropriate to censure others or are far scarcer than cases in which others are blameworthy (see Adams 1985: 21-24).

Since I accept ‘ought’ implies ‘can,’ I respond to the argument by rejecting the argument’s second premise, the view that ‘ought’ implies ‘can’ should be construed as ‘ought’ implies voluntarily ‘can.’ I consider and reject two reasons to accept this claim.

The first reason to think that ‘ought’ implies ‘can’ should be construed as ‘ought’ implies voluntarily ‘can’ has already been eluded to in the previous sub-section. On this view, the subject who flouts an all things considered obligation is blameworthy for doing so. So, if subjects can be blameworthy only for activities under voluntary control, then obligations can govern only those activities which are under their voluntary control. Thus, the most plausible variant of ‘ought’ implies ‘can’ is the view that ‘ought’ implies voluntarily ‘can.’

As I argued in §2.2.1.1, we should reject the view that subjects can be blameworthy only for those activities which are under their voluntary control. Rather, there is reason to think that subjects can be criticizable for activities which reflect on their desires, values, or cares, whether or not those activities are under voluntary control. Thus, considerations from blameworthiness suggest no reason to construe ‘ought’ implies ‘can’ as ‘ought’ implies voluntarily ‘can.’

The second reason to think that ‘ought’ implies ‘can’ should be construed as ‘ought’ implies voluntarily ‘can’ is the view that there is a conceptual entailment between being able to perform some activity and voluntarily performing that activity. On this view, if some subject can perform some activity, it is a conceptual truth that subject can voluntarily perform that activity.

As both Conor McHugh and the team of Philippe Chuard and Nicholas Southwood have argued, it is a mistake to think that if some subject can perform some activity, then that subject can voluntarily perform that activity (Chuard & Southwood 2009: 614–625, McHugh 2012: 86–93). Subjects can do all sorts of things which are not under their voluntary control. They can sneeze and can hiccup. They can fall asleep and can dream. They can hear sounds out their window and can see objects in their line of sight. They can tell whether an animal before them is a dog or a cat. They can recognize their loved ones on sight. Though these are all things subjects *can* do, none of these are things subjects *voluntarily* do.¹⁵ The fact that some subject can perform some activity is no guarantee that that subject can voluntarily perform that activity. So, ‘ought’ implies ‘can’ does not conceptually entail ‘ought’ implies *voluntarily* ‘can.’

Having considered and rejected two reasons to think that ‘ought’ implies ‘can’ should be construed as ‘ought’ implies voluntarily ‘can,’ I conclude that it is possible to endorse ‘ought’ implies ‘can’—as I myself do—without committing to the claim that beliefs are under voluntary control.

¹⁵ The examples of sneezing and falling asleep are from McHugh (2012: 85, 87). The examples of object discrimination and perception are from Chuard & Southwood (2009: 618).

2.2.2. *The Objection from the View that Obligations Must Guide Action*

A different objection to the norm of revision—and to norms of rationality more broadly—is due to Kathrin Glüer and Åsa Wikforss, who argue as follows: If norms of rationality are prescriptive, then such norms must make a difference in the guidance of action. Specifically, beliefs must be guided by those norms and not merely in conformity with them. But, it is implausible that beliefs are ever guided by norms, even when beliefs do conform to them. Therefore, it is implausible that there are prescriptive norms of rationality (Glüer & Wikforss 2013a).

As it turns out, I do not wish to claim that the norm of revision—or norms of rationality more generally—regulate beliefs. That is to say, I am not committed to the claim that the norm of revision is *prescriptive*, in the sense in which Glüer and Wikforss use that term. Rather, I am amenable to Glüer and Wikforss' proposal that the rationality of beliefs is merely a matter of conforming to certain norms, even if those beliefs are not regulated by those norms. A norm is a measuring tape, not a whip.

While the norm of revision is consistent with the Glüer and Wikforss claim that there are no prescriptive norms of rationality, one might exploit some of the points made by Glüer and Wikforss to develop a different argument, one which threatens the coherence of my claim that the norm of revision is not prescriptive. Specifically, one might argue as follows: rational obligations, if any, regulate beliefs. But, given some of the reasons Glüer and Wikforss adduce, it is implausible that norms of rationality do—or even, *could*—regulate beliefs. So, there are no norms of rationality.

I reject the first premise of this argument, the view that norms of rationality, if any, regulate belief. On the face of it, this might seem a strange move. For what would be the point of positing such norms if they do not—perhaps, *could* not—regulate belief? I would suggest that norms of rationality earn their keep in others ways. For instance, they serve as a useful theoretical tool for making sense of rational blameworthiness. What it is to say that someone is rationally blameworthy is to say that her beliefs do not conform to some all things considered obligation of rationality. Such an explanation can both clarify the nature of rational blameworthiness and help us identify instances of its kind. This is so even if this explanation is not reductive, i.e., even if rational obligation explains rational blameworthiness as much as rational blameworthiness explains rational obligation.

In this section, I have considered two objections to the view that there are norms of rationality and have argued that neither of these objections succeeds. In light of this result and in light of the intuitions which motivate the norm of revision, I conclude that we should accept the norm of revision.

I turn to defending the argument's second premise, which is a variant of 'ought' implies 'can.'

2.3 Epistemic 'Ought' Implies A Robust Psychological 'Can'

The argument's second premise is the claim that if some subject is *prima facie* rationally required to revise some mental state in response to certain evidence then she can revise that mental state. In particular, she can revise that mental state, given the psychological mechanisms and skills she already has.

Notice that it is not enough for a subject to have a robust psychological ability to revise some mental state that she could revise that state, were she to acquire additional psychological mechanisms or skills; she must already have any abilities or skills needed for the task. As previously argued, this subject need not be able to voluntarily bring about the relevant change. Rather, this subject might revise her mental state in the same way she falls asleep or recognizes her loved one's on sight—i.e., non-voluntarily.

In this section, I make two separable claims: first, I defend the argument's second premise, the claim that rational obligations are constrained by one's current psychological mechanisms and skills. Second, I argue against a different version of 'ought' implies 'can,' one on which one's rational obligations are constrained by one's current mental states. The latter result will be useful in defending the preferred variant of 'ought' implies 'can' from a certain objection.

Psychological mechanisms and skills include both idiosyncratic and more widely held capabilities. They also span the divide between learned and innate traits. Examples include: perceptual capacities, the ability to empathize with others, the ability to read Mandarin, the ability to add sums, the ability to discriminate between heard pitches, the ability to perceptually discriminate between dogs and cats, the ability to see movement, and the tendency to resolve overtly dissonant attitudes in the direction of greater coherence.

In order for a subject to currently have some psychological mechanism or skill, that subject must exhibit whatever trait or activity is relevant to that mechanism or skill in a suitably wide range of psychologically similar circumstances, i.e., in circumstances in which her psychology is extremely similar to its present state. For instance, in order for a subject to have the ability to speak Mandarin, it is not enough that she *might*, by some freak coincidence, speak the exact Mandarin words which convey her intended meaning. She must be able to produce correct Mandarin with some reliability, which this 'fluke-ish' ability doesn't permit. Likewise, it is not enough for some subject to (currently) have the ability to speak Mandarin that she *could* enroll in some courses and in this way come to learn the language. If this subject can speak Mandarin, she must reliably produce Mandarin in circumstances in which her psychology is very similar to its present state.

Circumstances in which this subject has studied Mandarin are circumstances in which her psychology is quite different from its present state and thus, these circumstances cannot make it the case that this subject is currently able to speak Mandarin.¹⁶

Psychological mechanisms and skills form a natural contrast with mental states, such as thoughts, moods, desires, sensations, and perceptual experiences. The belief that it is raining out is a not a mechanism or a skill, though it may rely on or be governed by mechanisms or skills. The same can be said of a feeling of exuberance, a searing pain in one's leg, a desire for pistachio ice cream, and a distracted mood. This ostensive characterization of mechanisms and skills on the one hand, and of mental states, on the other, serves present purposes.

The relevant 'ought' implies 'can' principle is motivated by intuitions about certain cases. For instance, someone who is fluent in English but who has only a minimal grasp of Mandarin cannot be rationally required to revise her inaccurate belief about what some text means just by looking at its Mandarin translation. In certain circumstances, someone who reads Mandarin fluently might be so required. Likewise, someone who lacks the ability to identify pitches by their sound cannot be rationally required to revise her belief about which key a certain song is in. Someone who has this ability—what is sometimes called *perfect pitch*—might be so required.

The question arises: why aren't non-readers of Mandarin rationally required to revise their assessments on the basis of Mandarin-encoded information, when their Mandarin-fluent counter-parts might be so required? And why aren't those who lack perfect pitch rationally required to revise their beliefs about which key a song is in, when those of their counter-parts blessed with perfect pitch might be so required? If we accept (2), and accept that rational obligations are constrained by psychological ability, we enjoy a straightforward explanation: these asymmetries are explained by differences in the relevant subjects' current repertoire of psychological abilities.

Reflection on these cases also makes it clear that rational obligation is constrained by one's current abilities, not (merely) by those abilities one might acquire. For, if rational obligation were constrained merely by those obligations one might acquire, we would have the intuition that the person who doesn't know Mandarin *is* rationally required to revise her belief about a Mandarin-encoded text. For, this person *might* learn Mandarin and so, Mandarin is among those abilities she might acquire.

Thus far, I've argued that rational obligation is constrained by the psychological abilities and skills one already possesses. The question arises whether rational obligation is also constrained by one's current mental states. That is, if you are unable, given your current mental states, to regulate some belief, does that fact preclude that you have a

¹⁶ I thank {*removed for anonymous review*} for the example of speaking Mandarin as a fluke.

rational obligation to revise that belief? I find that it does not. Suppose that you incorrectly believe that $8 \times 6 = 46$. While splitting a bill, you rely on this belief and as a result, come to an inaccurate view about how the bill should be split. The question is, rationally speaking, are you required to revise your belief about how the bill should be split? Let's suppose that, holding fixed your belief that $8 \times 6 = 46$, you cannot revise your belief about the bill. It seems to me that you are rationally required to revise your belief, your troublesome belief notwithstanding.

Or, to take another case, suppose that you are a juror in a criminal trial, one in which the evidence strongly suggests that the defendant is innocent of the crime of which he is accused. Finding the defendant to be a repugnant person, you develop a strong dislike of him. As a result of your antipathy, you focus on the most incriminating evidence, resulting in a belief that the defendant is guilty. Rationally speaking, should you revise your belief? I think the answer is yes, even if, holding fixed your antipathy for the defendant, you cannot revise your belief.

From the preceding cases, I conclude that you can be rationally required to revise a certain belief even if, holding fixed your other mental states, you are not psychologically able to revise that belief. Thus, while rational obligation is constrained by one's current psychological abilities and mechanisms, rational obligation is not constrained by one's current mental states.

Once it is clear that the defended version of 'ought' implies 'can' does not extend to the view that rational obligations are constrained by one's current mental states, it is possible to defend the view against a natural objection to epistemic 'ought' implies 'can.' Sharon Ryan articulates this objection as follows: suppose you have a much-loved friend who is accused of stealing from you. You believe your friend to be innocent, despite the overwhelming evidence against her. In fact, your love for your friend is so strong that you are psychologically incapable of revising your belief in your friend's innocence. According to Ryan, you are rationally obligated to revise your belief in light of the evidence, even if you are psychologically incapable of doing so. Ryan takes this case to refute the claim that epistemic 'ought' implies 'can' (Ryan 2003: 59).¹⁷

While Ryan's case threatens traditional versions of 'ought' implies 'can,' it leaves unscathed the variant I have proposed. For presumably, what explains your inability to revise your belief in your friend's innocence is your affection for your friend. But affection is a mental state, not a psychological mechanism or skill. On the variant of 'ought' implies 'can' proposed, rational obligation is constrained by one's current psychological mechanisms and skills, but not by one's current mental states. So, it is not a counter-example

¹⁷ See Graham (2011), Mizrahi (2012) and Ryan (2003) for other criticisms of 'ought' implies 'can.' For defenses, see Littlejohn (2012) and Vranas (2007).

to this version of ‘ought’ implies ‘can’ that you have a rational obligation which your affection makes you unable to fulfill.¹⁸

I have been supposing a certain understanding of Ryan’s case, one on which what explains your inability to revise your belief in your friend’s innocence is your affection for your friend. On this view, were you to somehow lose your affection for your friend, you would find yourself able to revise your belief in her innocence. If this is not the right interpretation—if it is rather that your psychological mechanisms are somehow inadequate to the task of revising your belief—then the ‘ought’ implies ‘can’ under consideration would predict that you have no obligation to revise your belief. But in *this* case, in which your current repertoire of psychological mechanisms and skills precludes the revision of your belief, it seems implausible that you might have an obligation to do so. This would be akin to claiming that a non-Mandarin speaker might have a rational obligation to revise her belief about the meaning of a Mandarin-encoded text. This claim is highly counter-intuitive.¹⁹

2.4 *The Argument Completed*

The conclusion of the argument is that for any subject and for any of that subject’s beliefs, that subject has a robust psychological capacity to revise that belief in response to counter-evidence. In particular, that subject can, holding fixed the psychological skills and mechanisms she already possesses, revise that belief. This ability must be something the subject herself possesses, but it need not be under her voluntary control.

3 An Objection to the Revisability View: *Idée Fixes*

While the argument from the norm of revision suggests a powerful reason for accepting the revisability view, the view makes certain predictions which may seem counter-intuitive. For, the view excludes from the class of belief any mental state that is not capable of being rationally revised, even if that mental state: is sincerely endorsed by its subject, guides its subject’s actions, and serves as a premise in a wide range of its subject’s inferences.

¹⁸ {removed for anonymous review} has pointed out to me that that this reading of the case suggests that you have an epistemic reason to abandon your affection for your friend, thus making the case one in which epistemic requirements impinge on the requirements of friendship. I don’t think this makes the case unusual. Surely epistemic requirements impinge not only on obligations of friendship, but also on obligations of love, of familial bonds, and of community ties. Such relationships are bound to cloud one’s judgment in ways that are likely to be epistemically pernicious. I don’t find this result alarming, since I think norms generated by such relationships trump epistemic norms.

¹⁹ See also McHugh (2012: footnote 62) and Vranas (2007: 183–84).

Consider the following case: Sheryl forms the view that her neighborhood's farmer's market takes place on Fridays. This view is formed in response to excellent evidence. But some point after forming this view, Sheryl suffers a minor brain lesion which leaves her cognitive faculties entirely intact except for the curious result that she cannot revise her view about the local farmer's market. She sees flyers advertising that the market has been rescheduled for Sundays, her friends repeatedly tell her the market is now on Sundays, she even visits the farmer's market on Sundays (quite by accident, since she doesn't anticipate its being held then), but she simply cannot revise her view that the market is held on Fridays. She shows up at the usual spot every Friday, bags and shopping list in tow. She tells anyone who asks that the market is on Fridays. She relies on this claim as a premise in a wide range of inferences, such as: *The market is on Fridays. Today is Friday. The market is today.* The lesion has transformed Sheryl's prosaic belief into an *idée fixe*. On the revisability view, this mental state is not a belief, even though it was formed in response to good evidence. It must instead be some other cognitive attitude, such as an entertained thought, an assumption, or a non-doxastic delusion.

The fact that the revisability view excludes *idées fixes* from the class of belief might suggest that the view is too strong. After all, Sheryl sincerely endorses this view, is guided by it in action, and relies on it in a wide range of inferences. If the revisability view entails that Sheryl's view and others like it do not count as beliefs, perhaps we should reject the revisability view, in favor of the following view:

THE ANTI-REVISABILITY VIEW OF BELIEF

There exists at least some subject such that: that subject has some belief p and has sufficiently strong, undefeated evidence that that *not- p* , and that subject is not able, holding fixed the psychological mechanisms and skills which she already possesses, to revise that belief.

3.1 Sincere Assertion, Action Guidance, and Inferential Promiscuity Aren't Sufficient for Belief

In this section, my aim is to undermine the intuition that *idées fixes* are beliefs. My strategy is to challenge the theory which I take to drive the intuition. In this way, I take aim at both the intuition and at the attractive theory which underpins it. The theory in question is the view that any mental state which is sincerely asserted, action guiding, and inferentially promiscuous is necessarily a belief. I will cast doubt on this view by suggesting that there are counter-examples to it. I will first consider whether these traits are individually sufficient for belief and will then consider whether these traits are jointly sufficient for belief.

First, consider sincere assertion. It might be that as a general rule, if some subject sincerely asserts p , then that subject believes p . But, as multiple theorists have argued,

this rule has exceptions, most famously in cases of self-deception. Consider the subject who believes that his husband is cheating on him but who also vociferously denies—to both himself and others—that this is the case. Such a subject might sincerely assert that his husband is faithful, while nevertheless exhibiting behavior that is consistent with his belief, such as feeling sad when his husband calls yet again to say that he will be working late, or asking his husband more questions than usual about his whereabouts. If this sort of case is so much as possible, the fact that someone sincerely asserts p does not guarantee that that person believes p (Cohen 1992: 68-73, Mandelbaum 2014: 79-81).

Second, consider action guidance. It might be thought that if a mental state guides action in the right way, then this state is necessarily a belief. However, as many theorists have pointed out, mental states other than belief, including pretenses and suppositions, can guide action (Gendler 2007, 2008; Velleman 2000). For instance, suppose you are pretending to be an elephant. You might wave your trunk and walk clumsily and slowly. You needn't believe you are an elephant in order for your pretense to govern your actions in this way (Velleman 2000).

In response to the suggestion that mental states such as pretense can guide action, it might be claimed that these states do not motivate actions across a wide range of contexts, whereas beliefs do. The problem with this reply is that it seems possible that a subject might have a chronic, far-reaching fantasy of the sort which guides her actions in a wide range of circumstances. (I will describe such a case shortly and will argue that the fantasy in question is not a belief.)

Third, consider inferential promiscuity, which is availability as a premise in a wide range of inferences. Mental states which are inferentially promiscuous have a specific kind of productive power; they can inferentially-causally generate new mental states.

Inferential promiscuity might be a necessary condition on belief, but there is reason doubt that it is sufficient for belief. For notice that suppositions also exhibit inferential promiscuity. Consider the geometry student tasked with proving that no triangle has four sides. To do this, this student might suppose, for the sake of proving otherwise, that there is a triangle which has four sides and then attempt to generate a contradiction, as part of a demonstration that the original supposition is false. We might say that this student 'hypothetically adds to her stock of beliefs' that there is a triangle which has four sides, but whatever it is to hypothetically add to one's stock of beliefs that there is a triangle which has four sides, it is not to *believe* that there is a triangle that has four sides. It is to suppose it or merely entertain it.

One objection to the view that suppositions are inferentially promiscuous comes from the claim that suppositions cannot contribute to a *wide* range of psychological in-

ferences, whereas beliefs can.²⁰ For, while suppositions can be used as premises in inferences by *reductio*, they cannot serve as premises in other inferences. In contrast, beliefs can serve as premises in most any psychological inference.²¹

Though initially an appealing thought, it is false that the inferential power of supposition is more circumscribed than that of belief. The grain of truth in this claim is that we ordinarily do not make suppositions just for the sake of drawing out their entailments. However, we *might* do this, as we sometimes do in logic and philosophy (and, one might argue, mathematics). Suppositions have at least as much inferential power as beliefs. For, whatever can be believed can be supposed. And whatever inference can be drawn from a set of beliefs can be drawn from an analogous set of suppositions. For instance, I can believe *p* and believe *if p, then q* and thereby come to believe *q*. Likewise, I can suppose *p* and suppose *if p, then q* and thereby come to suppose *q*.

While it is false that the inferential promiscuity of suppositions is more circumscribed than that of belief, there is a different trait—one in the vicinity of inferential promiscuity—which promises to distinguish supposition from belief. This is the trait of serving as a premise in a wide range of psychological inferences, whose conclusions are *believed* (and not merely supposed). While beliefs likely have this trait, suppositions probably do not, since the conclusions of inferences involving suppositions are usually supposed, not believed. For instance, if I suppose, for the sake of argument *if cheese is blue, then Argentina is in Europe*, and I also suppose *cheese is blue*, then I might come to suppose *Argentina is in Europe*, but I will probably not believe that Argentina is in Europe.

Thus, one might suggest a new sufficient condition on belief, along the following lines: If some mental state can serve in a wide range of inferences, whose conclusions are believed, then that state is a belief. This criterion is quite attractive. But notice that it does not suggest a reason to think that *idées fixes* are beliefs. For while it was relatively clear that *idées fixes* are inferentially promiscuous, it is much less clear whether the conclusions of inferences in which *idées fixes* figure are believed.

Suppose, for instance, that Sheryl puts her *idée fixe* to use in the following inference: *The farmer's market is on Friday. Friday is the day after Thursday. The farmer's market is the day after Thursday*. Does Sheryl believe that the farmer's market takes place the day after Thursday? Or does she merely hold this view in a non-doxastic way, connected as it is to her *idée fixe*? I think the answer is extremely unclear. Thus, the proposed condition on belief does not seem particularly useful as a route of arguing that *idées fixes* are beliefs, even if this condition is sufficient for belief.

²⁰ I thank {*removed for anonymous review*} for raising this concern.

²¹ Or rather, beliefs can serve a premise-*like* role in psychological inferences. Premises are abstract entities, connected as they are to arguments, which are also abstract.

So far, I've attempted to cast doubt on the view that sincere assertion, action guidance, and inferential promiscuity are individually sufficient for belief. However, it might be claimed that even if these traits do not individually suffice for belief, they nevertheless *jointly* suffice for belief: on this view, any state which has all three traits is necessarily a belief, its unrevisability notwithstanding.²²

Surely, any mental state which is sincerely asserted, action guiding, and inferentially promiscuous is very *likely* to be a belief. However, I think it is false that any such state is necessarily a belief. To make this point, I draw on a case in which a subject has a chronic fantasy, one which is both in conflict with and yet fueled by what she believes. I will argue that this fantasy satisfies all three traits and yet, is not a belief.

Consider Carolyn, the teenaged protagonist of the novel *Uncle Vampire*, who develops a detailed and chronic fantasy that her uncle is a vampire. Carolyn sincerely endorses this fantasy, acts on it in a wide variety of circumstances, and relies on it in a wide range of inferences. The fantasy plays a pervasive and important role in her life. In a moment of extreme distress, Carolyn reveals to a therapist that her uncle molests her. It becomes clear that Carolyn's fantasy is some part of her attempt to deal with a traumatic reality.

The question is, does Carolyn *believe* that her uncle is a vampire, or does she represent this view in some other way, as in a non-doxastic pretense or delusion? There are at least two reasons to think that Carolyn does not believe that her uncle is a vampire. First, apart from her 'vampire' fantasy, Carolyn is a bright and highly sociable American teenager. While it is certainly coherent to suggest that Carolyn believes that vampires exist, it seems rather unlikely that a person of Carolyn's developmental age and cultural context should harbor such a belief.

Second, on at least one interpretation of Carolyn's psychology, her 'vampire' fantasy is in conflict with what she believes. For, Carolyn's fantasy represents that her uncle is a vampire (and not a molester), whereas her belief represents that her uncle is a molester. There is reason to think that beliefs can and do conflict, but it would be odd to claim that some belief can sustain another belief with which it conflicts. The view that Carolyn's fantasy is a belief would involve this surprising relationship between conflicting beliefs.²³

²² I thank {*removed for anonymous review*} for pressing me on this point.

²³ It might be claimed that Carolyn does not come to believe that her uncle is a molester until the disclosure in the therapist's office. I think this is implausible because it is presumably Carolyn's awareness that she is being abused which explains why she develops the 'vampire' fantasy. In any case, we might describe the case in a way which makes it clear she has the belief before the disclosure.

Consider other plausible instances of conflicting belief. Perhaps you can believe that if you don't say 'cancer' out loud, your partner's biopsy will come back cancer-free, even whilst believing that it is false that there is any connection between your saying 'cancer' and your partner's having cancer (Huddleston 2012). Or, perhaps you can believe that there is no external world, having convinced yourself by philosophical argument of this result, whilst also believing—because you cannot help it—that an external world exists.²⁴ In both of these cases, what permits your beliefs to conflict is that they derive from different kinds of evidence.

Notice that Carolyn's states are unlike the paradigm instances of conflicting beliefs. Her belief that her uncle is a molester somehow sustains her fantasy that her uncle is a vampire (and not a molester). So, Carolyn's views lack the etiological distinctness which is the hallmark of other pairs of conflicting beliefs. While it is not incoherent that two conflicting beliefs might be causally connected, given that this structure has no precedent, it would seem *ad hoc* to suppose that Carolyn's fantasy is a belief. Less *ad hoc* would be to suppose that Carolyn's fantasy is a delusion, a pretense, or some other sort of non-doxastic state. If this is right, then Carolyn's fantasy is a counter-example to the view that sincere assertion, action guidance, and inferential promiscuity are jointly sufficient for belief.

I conclude that we have good reason to doubt that sincere assertion, action guidance, and inferential promiscuity are sufficient for belief. So, we should be at least doubtful whether *idées fixes* are beliefs. Thus, the fact that the revisability view excludes such states from the class of belief does not count as a reason to reject the revisability view.

Moreover, the argument in favor of the revisability view constitutes a reason to exclude *idées fixes* from the class of belief: because the norm of revision and 'ought' implies 'can' cannot be simultaneously true of these states, we should not class them as beliefs.

Consider again Sheryl's *idée fixe* that the farmer's market is on Fridays. Notice that the following two claims—which are instances of the first two premises of the argument from the norm of revision—cannot be simultaneously true of Sheryl:

- (7) If Sheryl has sufficiently strong, undefeated evidence that the farmer's market is not on Friday, then Sheryl has a *prima facie* obligation of rationality to revise her *idée fixe*.
- (8) If Sheryl has a *prima facie* obligation of rationality to revise her *idée fixe*, then, holding fixed those psychological mechanisms and skills Sheryl currently possesses, Sheryl can revise her *idée fixe*.

²⁴ This example is from {*removed for anonymous review*.}

On the assumption that Sheryl has sufficiently strong, undefeated evidence that the farmer's market is not on Friday, Sheryl's *idée fixe* cannot simultaneously satisfy (7) and (8) and yet, beliefs must satisfy (7) and (8). For, if Sheryl's *idée fixe* is a belief, then (7) says Sheryl is obligated to revise her *idée fixe*. But, if Sheryl is obligated to revise her *idée fixe*, then (8) tells us that Sheryl must have a robust psychological ability to revise that *idée fixe*. The fact that Sheryl's *idée fixe* cannot simultaneously satisfy both of these claims counts as a strong reason to exclude it—and states like it—from the class of belief.

In short, because *idées fixes* exhibit belief-like traits, such as being sincerely endorsed, action guiding, and inferentially promiscuous, we mistook them for beliefs. But closer reflection reveals that *idées fixes* are not beliefs, but are rather pretenses, assumptions, entertained thoughts, or some other non-doxastic state.

4 Summary

I have defended the view that subjects necessarily have a robust psychological capacity to revise their beliefs in response to counter-evidence. This view can permit into the class of belief mental states which are never contravened by the evidence. The view can also permit into the class of belief many irrational states, such as those sustained by confirmation bias. At the same time, the view has the surprising result that *idées fixes* are not beliefs. Though initially counter-intuitive, I have argued that this result should in fact be welcomed, as there are reasons to reject the view which undergirds the intuition that *idées fixes* are beliefs.

Along the way, I have defended the existence of doxastic obligations. I have also developed a novel variant of epistemic 'ought' implies 'can' and have argued that this version outperforms traditional versions of the view. Finally, I have shown how the character-based approach to moral blameworthiness might be extended to the epistemic domain, to help make sense of epistemic blameworthiness.²⁵

REFERENCES

- Adams, R. M. (1985). Involuntary sins. *The Philosophical Review*, 94(1), 3-31.
- Adler, J. (2002). *Beliefs wwn ethics*. Cambridge, MA: MIT Press.
- Arpaly, N. (2004). *Unprincipled virtue: An inquiry into moral agency*. Oxford: Oxford University Press.
- Alston, W. P. (1988). The deontological conception of epistemic justification. *Philosophical Perspectives*, 2, 257-299.

²⁵ I am extremely grateful to the following people for their feedback: {removed for anonymous review}.

- Bayne, T. & Pacherie, E. (2005). In defence of the doxastic conception of delusions.' *Mind & Language*, 20(2): 163-188.
- Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford: Oxford University Press.
- Boudry, M. & Coyne, J. (2016). Disbelief in belief: On the cognitive status of supernatural beliefs. *Philosophical Psychology*, 29(4), 601-615.
- Broome, J. (2007). Wide or narrow scope? *Mind*, 116(462), 359-370.
- Brownstein, M. (2016). Attributionism and moral responsibility for implicit bias. *Review of Philosophy and Psychology*, 7(4), 765-786.
- Card, C. (1988). Gratitude and obligation. *American Philosophical Quarterly*, 25(2), 115-127.
- Chuard, P. & Southwood, N. (2009). Epistemic norms without voluntary control. *Noûs*, 43(4), 599-632.
- Cohen, L. J. (1992). *An essay on belief and acceptance*. Oxford: Clarendon Press.
- Currie, G. & Ravenscroft, I. (2002). *Recreative minds*. Oxford: Oxford University Press.
- Currie, G. & Jureidini, J. (2001). Delusion, rationality, empathy: commentary on Martin Davies et al. *Philosophy, Psychiatry, & Psychology*, 8(2), 159-162.
- Davidson, D. (1984). *Inquiries into truth and interpretation*. Clarendon: Oxford University Press.
- Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT Press.
- Egan, A. (2009). Imagination, delusion, and self-deception, in *Delusion and Self-deception: Affective and Motivational Influences on Belief Formation*, eds. T. Bayne and J. Fernandez. New York: Psychology Press.
- Frankish, K. (2012). Delusions, levels of belief, and non-doxastic acceptances,' *Neuroethics* 5(1): 23-27.
- Friedman, J. (2013). Suspended Judgment. *Philosophical Studies* 162(2): 165-181.
- Gendler, T. S. (2007). Self-deception as Pretense, *Philosophical Perspectives* 21(1): 231-258.
- Gendler, T. S. (2008). Alief and belief, *The Journal of Philosophy* 105(10): 634-663.
- Gertler, B. (2011). Self-knowledge and the transparency of belief. *Self-knowledge*, 125-145.
- Glüer, K. & Wikforss, Å. (2013a). Against belief normativity. In *The Aim of Belief*, ed. T. Chan, 80-99. Oxford: Oxford University Press.
- Glüer, K. & Wikforss, Å. (2013b). Aiming at truth: on the role of belief. *Teorema* 32(3): 137-162.
- Graham, P. A. (2011). 'Ought' and ability. *The Philosophical Review* 120: 337- 382.
- Grant, C. D. (2015). *Uncle Vampire*. Open Road Media.

- Harmon-Jones, E. (2000). A cognitive dissonance theory perspective on the role of emotion in the maintenance and change of beliefs and attitudes, *Emotions and Beliefs* 185-211.
- Harmon-Jones, E. & Harmon-Jones, C. (2007). Cognitive dissonance theory after 50 years of development.
- Hieronymi, P. (2008). Responsibility for believing. *Synthese* 161: 357–373.
- Hieronymi, P. (2014). Reflection and responsibility. *Philosophy & Public Affairs*, 42(1), 3-41.
- Huddleston, A. (2012). 'Naughty Beliefs.' *Philosophical Studies* 160: 209-22.
- Jaworska, A. (1999). Respecting the margins of agency: Alzheimer's patients and the capacity to value. *Philosophy & public affairs*, 28(2), 105-138.
- Jaworska, A. (2007). Caring and internality. *Philosophy and Phenomenological Research*, 74(3), 529-568.
- King, A. (2014). Actions that we ought, but can't. *Ratio*, 27(3), 316-327.
- Kolodny, N. (2005). Why be rational? *Mind*, 114(455), 509-563.
- Levy, N. (2015). Neither fish nor fowl: Implicit attitudes as patchy endorsements. *Noûs*, 49(4), 800-823.
- Levy, N. (2017). Religious beliefs are factual beliefs: Content does not correlate with context sensitivity. *Cognition*, 161, 109-116.
- Littlejohn, C. (2012). Does 'ought' still imply 'can'? *Philosophia* 40(4): 821-828.
- Madva, A. (2015). Why implicit attitudes are (probably) not beliefs. *Synthese*, 1-26.
- Madva, A. & Brownstein, M. (2016). Stereotypes, prejudice, and the taxonomy of the implicit social mind. *Noûs*.
- Mandelbaum, E. (2014). Thinking is believing. *Inquiry* 57(1): 55-96.
- Mandelbaum, E. (2016). Attitude, inference, association: On the propositional structure of implicit bias. *Noûs*.
- McHugh, C. (2012). Epistemic deontology and voluntariness. *Erkenntnis*, 77(1), 65-94.
- McHugh, C. (2013). Epistemic responsibility and doxastic agency. *Philosophical Issues*, 23(1), 132-157.
- McHugh, C. (2017). Attitudinal control. *Synthese*, 194(8), 2745-2762.
- Mizrahi, M. (2012). Does 'ought' imply 'can' from an epistemic point of view?' *Philosophia* 40: 829-840.
- Mullen, R., & Gillett, G. (2014). Delusions: A different kind of belief? *Philosophy, Psychiatry, & Psychology*, 21(1), 27-37.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2(2): 175.
- Radden, J. (2010). *On Delusion*. Routledge.

- Radden, J. (2014). Belief as delusional and delusion as belief. *Philosophy, Psychiatry, & Psychology*, 21(1), 43-46.
- Ryan, S. (2003). Doxastic compatibilism and the ethics of belief. *Philosophical Studies* 114(1): 47-79.
- Rusin, J. (2010). Reconciling epistemic duty with non-voluntary belief. *International Journal of the Humanities*, 8(8).
- Schwind, C., Buder, J., Cress, U. & Hesse, F. W. (2012). Preference-inconsistent recommendations: An effective approach for reducing confirmation bias and stimulating divergent thinking? *Computers & Education* 58(2): 787-796.
- Schwitzgebel, E. (2001). In-between believing. *The Philosophical Quarterly* 51(202): 76-82.
- Schwitzgebel, E. (2002). A dispositional, phenomenal account of belief. *Noûs* 36(2): 249-275.
- Schwitzgebel, E. (2010). Acting contrary to our professed beliefs or the gulf between occurrent judgment and dispositional belief. *Pacific Philosophical Quarterly*, 91(4), 531-553.
- Shah, N. & Velleman, J. D. (2005). Doxastic deliberation. *The Philosophical Review* 497-534.
- Sher, G. (2009). *Who knew? Responsibility without awareness*. Oxford: Oxford University Press.
- Shoemaker, D. W. (2003). Caring, identification, and agency. *Ethics*, 114(1), 88-118.
- Shoemaker, D. (2011). Attributability, answerability, and accountability: Toward a wider theory of moral responsibility. *Ethics*, 121(3), 602-632.
- Smith, A. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics* 115(2): 236-271.
- Smith, A. (2008). Control, responsibility, and moral assessment. *Philosophical Studies* 138: 367-392.
- Smith, A. (2012). Attributability, answerability, and accountability: In defense of a unified account. *Ethics* 122(3): 575-589.
- Sripada, C.S. (2010). The deep self model and asymmetries in folk judgments about intentional action. *Philosophical Studies* 151(2): 159-176.
- Sripada, C. (2015). Self-expression: A deep self theory of moral responsibility. *Philosophical Studies*.
- Steup, M. (2008). Doxastic freedom. *Synthese* 161(3): 375-392.
- Stich, S. P. (1978). Beliefs and subdoxastic states. *Philosophy of Science* 499-518.
- Stephens, G. L. & Graham, G. (2004). Reconceiving delusion. *International Review of Psychiatry*, 16(3), 236-241.
- Viedge, N. (2016). Defending evidence-resistant beliefs. *Pacific Philosophical Quarterly*.

- Van Leeuwen, N. (2014). Religious credence is not factual belief. *Cognition*, 133(3), 698-715.
- Van Leeuwen, N. (2016). Beyond fakers and fanatics: A reply. *Philosophical Psychology*, 29(4), 616-621.
- Van Leeuwen, N. (2017). Two paradigms for religious representation: The physicist and the playground (a reply to Levy). *Cognition*, 164, 206-211.
- Velleman, D. (2000). *The Possibility of Practical Reason*. Oxford: Oxford University Press.
- Vranas, P. (2007). I ought, therefore I can. *Philosophical Studies* 136: 167 - 216.
- Watson, G. (1975). Free agency. *Journal of Philosophy* 72: 205-220.
- Watson, G. (1996). Two faces of responsibility. *Philosophical Topics* 24(2): 227-248.