BELIEF'S MINIMAL RATIONALITY

MARIANNA BERGAMASCHI GANAPINI

Many of our beliefs behave irrationally: this is hardly news to anyone. Although beliefs' irrational tendencies need to be taken into account, this paper argues that beliefs necessarily preserve at least a minimal level of rationality. This view offers a plausible picture of what makes belief unique and will help us to set beliefs apart from other cognitive attitudes (e.g., imagination, acceptance).

In philosophy and cognitive science, mental attitude types (e.g., imagination, belief, and desire) are often defined in terms of their input and output (Fodor, 1985; Nichols and Stich, 2003). When applied to belief, it has long been argued that – on the output side – belief is inferentially promiscuous and action-guiding. On the input side, evidence is usually that which motivates one to form and revise one's beliefs. This means that for an attitude to count as a belief it must be prone to react to the evidence establishing the truth of its content, as well as to any counter-evidence disproving it. These core tendencies are also usually linked to belief's standards of doxastic rationality. Piggybacking on this mainstream view, which I call Traditionalism, many philosophers maintain that beliefs are *in fact* expected to behave rationally. This is Strong Traditionalism. Contrary to the *strong* version of Traditionalism, however, a group of Revisionist philosophers have recently argued that it is not *necessary* for a belief to have a significant and widespread impact on our theoretical and practical reasoning, and/or to appropriately respond to the relevant evidence. On a popular version

of this view, inferentially patchy and behaviorally inert attitudes can still count as beliefs so long as they are expressed through sincere assertions.

In this paper, after showing that Revisionism faces a number of objections, I propose an alternative, midway position which I call *Moderate* Traditionalism. Although irrational, inert, and patchy beliefs are possible, I argue that belief's key marker is exemplified in its reaction to (detected) irrationality. This is belief's *minimal* rationality. This view has some clear advantages over its immediate competitors. With respect to Strong Traditionalism, Moderate Traditionalism fares better in addressing the evidence of human irrationality. In opposing Revisionism, Moderate Traditionalism offers a plausible picture of what makes belief unique. Since I am trying to carve out a convincing psychological theory, Moderate Traditionalism will also allow me to offer a clear way to show that belief is functionally different from other cognitive attitudes.

1. Traditionalism and Revisionism

It is common in philosophy to associate beliefs with their action-guiding role, the production of new attitudes, as well as the characteristic of responding to evidence. These roles are largely taken to reflect standards of epistemic, procedural, and practical rationality. Beliefs motivate actions in the sense that they tend to 'combine' with conative attitudes (e.g., desiring) to produce actions that satisfy those conative attitudes only when the beliefs in question are true (Davidson, 1963; Smith, 1994). Desires and beliefs also provide *rational* explanations for the actions they tend to produce. If I believe that it is raining and I wish to avoid getting wet, I will be motivated or disposed to carry an umbrella if I go out. This is the practically-rational thing to do for me, ceteris paribus. Belief's effects are not limited to

action: beliefs cause the production of other beliefs (Glüer and Wikforss, 2013a; Stich 1978: 507). For instance, believing that *if* the cat is sleeping then she is in the bedroom, while also believing that in fact the cat is sleeping means being disposed to believe that she can be found in the bedroom. Presumably that does not mean that we in fact believe everything that follows from our current beliefs. However, we are at least prone to endorse a proposition if we attend to the fact that that proposition follows from, or is coherent with, other things that we hold true. Again, on a popular reading of this, the ability to combine and produce new beliefs that are inferentially integrated in this way is a key feature of belief, that is, part of its very nature. Mental state types not only have defining downstream consequences: they are also typically sensitive to certain kinds of inputs. Belief is sensitive to things we perceive or believe that are relevant to its content. Not surprisingly, then, my belief that the cat is sleeping in the bedroom will be prone to change according to whether, for example, I see the cat playing in the kitchen instead. And this disposition is consistent with the rational norms that set the epistemic standards for how to form and modify beliefs.

This analysis follows from a longstanding view that we can call 'Traditionalism'. Traditionalists by and large maintain that it is a necessary condition for being a belief that an attitude is *set* to behave in a rational way.[1] It is notoriously hard to determine rational tendencies, as these may be masked by other factors. Thus, when we see suboptimal behavior, we may ask ourselves if it is a problem of performance or something running deeper than that. In response, some have adopted a stronger version of Traditionalism, which furthermore requires beliefs to *exhibit* their rational tendency:

[1] The idea that we are fundamentally rational is fairly widespread (for instance, see Davidson (1984), Dennett (1987) and those who espouse Bayesianism, e.g., Tenenbaum et al., (2011)).

Strong Traditionalism: mental attitude A is a belief only if A is mostly doxastically rational, in the sense that it is (i) sensitive to the relevant evidence, (ii) inferentially integrated with other beliefs and intentional attitudes, and it (iii) causes actions (when coupled with the appropriate conative attitude, in the right circumstances).[2]

The term 'rational' is used for any belief conforming to *standards* of rationality, whereas 'irrational' beliefs subvert those standards. Roughly, a belief is *epistemically* irrational if held in the face of insufficient, available evidence.[3] Also, a belief is *procedurally* irrational if it does not produce inferential effects to preserve coherence with other attitudes.[4] Finally, a belief is *practically* irrational if it does not produce behavioral/emotional effects, and does not enjoy the role of guiding our practices of practical reasoning.[5]

---

[2] See Adler (2002), Currie and Ravenscroft (2002), Gendler (2008), Rey (1988), Van Leeuwen (2014), to name a few. Van Leeuwen (2014), for instance, claims that religious credences are *not* (factual) beliefs because on his view (factual) beliefs have specific input-output conditions (e.g., evidence-sensitivity) that religious credences do not share. That said, it is possible that some of these authors may be willing to endorse a more moderate form of Traditionalism, similar to the Traditionalism I advocate for below.

[3] Available evidence is not necessarily the same thing as what agents *take* to be evidence since agents may be wrong about what constitutes evidence for what they believe.

[4] As intended here, procedural rationality encompasses the structural and substantive inferential effects usually associated with belief (see Scanlon 2007). Beliefs support the formation of other attitudes and they also stand in (formal and non-formal) structural relations with other beliefs. Some of the coherence standards I will mention in the last section fall within this sense of rationality.

[5] It is admittedly slightly odd to add this to the stock of *doxastic* rationalities, since it in fact requires one to take some action, or to react in a certain way, rather than forming a belief. However, since

It is often said that cognitive attitudes are those attitudes that present their contents as true (Velleman, 2000), and beliefs are paradigmatic examples of cognitive attitudes. Beliefs do not, however, exhaust cognitive attitudes. Among non-doxastic cognitive attitudes are imagination, assumption, acceptance, hypothesis, and the like, which we can call "secondary cognitive attitudes" (Van Leeuwen, 2009). On the Strong Traditionalist view, if a cognitive attitude fails to show rational belief-like behavior, then we must conclude that it is not a belief after all, but some neighboring cognitive attitude (e.g., imagination[6], acceptance[7], alief[8], quasi-belief[9], non-standard belief[10]).

Contrary to this view, however, it has become increasingly apparent that some of our beliefs – or what we would intuitively call 'beliefs' – are at times behaviorally inert, immune to evidence, or inferentially encapsulated. Whatever rational role belief has, whatever requirements belief is subject to, beliefs (and people more generally) can be bad reasoners,

---

beliefs are meant to be action-guiding, beliefs that do not motivate us are considered failures qua beliefs. Note that in all of these formulations I mostly follow Bortolotti (2010). Let me also point out that the argument of this paper does not stem from a philosophical analysis of rationality per se. I will thus not focus on trying to figure out the correct formulations of the standards of rationality or ask whether or not these standards constitute normative requirements (Broome, 2007; 2013).

[6] See Currie and Ravenscroft (2002), and Currie and Jureidini (2002), who argue that delusions are imaginings, not beliefs.

[7] Frankish (2004; 2011), Shah & Velleman (2005).

[8] Gendler (2008) suggests that some evidence-immune attitudes are not really beliefs but aliefs. See Glüer and Wikforss (2013b) and Mandelbaum (2013) for concerns regarding Gendler's proposal.

[9] Schwitzgebel (2002: 252) and Schellenberg (2013) argue that 'belief' is a vague term that stands on a continuum.

[10] In this last category, we find the notion of "assent" offered by de Sousa (1971), Dennett's "opinion" (1978), Rey's "avowal" (1988), Sperber's "reflective belief" (1996) and Van Leeuwen's "credences" (2014).

falling short of rational standards (Kahneman and Tversky, 1979; Stein, 1996). As a result, dissonant or poorly integrated beliefs may be more common than previously thought. This has led some to believe that our cognitive system is fragmented (e.g., Lewis 1982; Stalnaker, 1984; Greco, 2015), and that beliefs may have minimal global effects on our minds since they are 'stored' in disparate compartments (Mandelbaum, 2016). Numerous studies appear to show that humans suffer from "belief-perseverance": at times they refuse to change their mind even when they have been proven wrong (e.g., Anderson, 1995; Guenther & Alicke 2008: 706; Mercier and Sperber, 2011: 68; Kahn, 2016). Moreover, there is evidence that we automatically come to believe whatever proposition is presented to us, independent of the evidence (Mandelbaum, 2019). Things do not look better once we turn to belief's action-guiding role: even sincere assertions are sometimes poor indicators of what agents will do (Wicker, 1969: 75; Ajzen, 2005; Fabrigar et al., 2006).

The worry now is that this body of evidence forces Strong Traditionalism to say that we have few real beliefs owing to the fact that so few attitudes pass Traditionalism's "rationality constraint" (Bortolotti, 2010; Helton, 2018).[11] Yet one may rightly suggest that

---

[11] A word of caution is needed here. As recently pointed out by Van Leeuwen (2018), there is a tendency to focus on belief's irrationality and forget the large number of rational beliefs that serve us in the background. So, perhaps it is unfair to claim that Strong Traditionalism would leave us with too few beliefs, because many of our beliefs are in fact rational. It still seems true, however, that Strong Traditionalism forces us to exclude from the category 'belief' things that we are normally happy to call beliefs. In other words, the worry is that Strong Traditionalism is, so-to-speak, throwing the baby out with the bathwater by imposing a too-rigid rationality-constraint on belief. Therefore, a more inclusive and fine-grained view – such as the one I propose below – would be preferable to Strong Traditionalism. Thank you to an anonymous referee for pushing me to clarify this point.

it is a "desideratum" of "our theorizing about belief" that the view we endorse does not directly conflict with the plausible assumption that "humans have many beliefs" (Helton, 2018: 2):

Too Few Objection: if we impose rationality constraints on belief, then we will be forced to say that we only have very few beliefs.

As a reaction to this objection, a group of Revisionist philosophers have argued that the Traditionalist position is too strict, owing to which we should "loosen up" belief's ties to evidence, action, and reasoning. On this view, then, it is not necessary for beliefs to consistently motivate us in action and reasoning, nor that they ought to respond appropriately to evidence. As Bortolotti (2010: 172) explains,

> some beliefs have the capacity for guiding action, but action guidance
>
> should be listed among the desiderata for rational beliefs, and not among
>
> the necessary conditions for beliefs ascription.

In response to this, a move that is often made is that patchy, and behaviorally- or inferentially-fragmented, cognitive attitudes may still count as beliefs provided they are expressed through sincere assertions or if they 'feel' true in some sense (Zimmerman, 2007). Accordingly, we do not need to chain ourselves to a narrow conception of belief that inexorably ties belief to its rational effects, since sincere assertion is enough to track beliefs:

Revisionism: for attitude A to be a belief, it suffices that A is expressed through sincere assertion: it is not necessary that A is mostly doxastically rational (i.e., sensitive to evidence, action-guiding, or inferentially-integrated).

To be sure, the Revisionist position has an intuitive appeal. For starters, as we saw, there is strong evidence that humans are more irrational than we thought. So, against the Strong Traditionalist, the Revisionist raises the reasonable worry that if we were to expect beliefs to be consistently rational, we would have to exclude from the category 'belief' many cognitive attitudes that we are usually happy to call beliefs (Bortolotti, 2010). Another attractive aspect of Revisionism stems from the oft-cited argument that an agent's "view of the world" is reflected in their beliefs. Philosophers working on self-knowledge have argued that we can seamlessly move from the question "do I believe P?" to the factual question "is P true?", owing to which it seems we have privileged access to our beliefs (Moran, 2001; Boyle 2011). This appears to indicate that sincerely asserting that p and believing that p are basically the same thing, and that people seem to have first-personal authority on what they take to be true. What this means is that, if people introspect and sincerely assert something, then we should probably give them credit. To boot, it seems that Revisionism gets our folk-psychological intuitions right: recent studies show that across cultures and languages "behavioral integration" is not necessarily a key factor when people make belief-attributions: "assertions are first taken into account, and when an agent sincerely asserts that p, nonlinguistic behavioral evidence is disregarded" (Rose et al., 2017: 194). According to these

studies, our folk-psychological intuitions seem to go the Revisionist's way: assertion is the mark of belief.12

2. Against Revisionism

In this section of the paper, I will level two objections against Revisionism. The first objection is that Revisionism has a 'demarcation' problem. This threatens its ability to correctly identify the extension of the category 'belief'. The second objection is that the Revisionist puts too much faith in sincere assertion in the face of recent findings about people's lack of self-knowledge and their inability to introspect. I will consider these objections in turn.

2.1 Revisionism's Demarcation Failure: The Too Many Objection

Belief is a cognitive attitude but the world of *cognitive* attitudes is densely populated. To clarify, cognitive attitudes are those attitudes that present their contents as true, and belief is only one of them (Velleman, 2000). Imagination, assumption, acceptance, and other non-doxastic cognitive attitudes are often called *secondary* cognitive attitudes (Van Leeuwen, 2009; 2014). Belief and secondary cognitive attitudes function differently, and a

---

12 Since I will criticize Revisionism below, it is important to clarify now that in this paper I am *not* trying to fully capture the extension of the pre-theoretical use of the word 'belief'. Nor am I trying to say that lay people are in error in their use of the word 'belief'. What I am doing here is appealing to a mature philosophy of mind to carve out importantly different mental states by looking at their functional role. Although, all things being equal, it would be preferable to have a theoretical view of 'belief' that squares with our pre-theoretical intuitions, *fully* conforming to our folk use of 'belief' is not a prerequisite of doing good psychology.

mature philosophy of mind needs to tell us how to distinguish them. And this raises the following worry for the Revisionist:

The Too Many Objection: if a cognitive attitude can count as a belief even if it is all-over irrational then we have no way to clearly demarcate beliefs from secondary attitudes. As a result, we may end up inflating the category 'belief'.

There are two aspects to the Too Many Objection. The first aspect has to do with Revisionism's inability to explain how belief is functionally distinct from other cognitive attitudes. The second element to consider is that, as a result of the demarcation problem, Revisionism has an extensional problem and ends up miscategorizing cognitive attitudes.

Let us focus first on the demarcation problem. Revisionism endorses what you might call a 'mixed-bag view of belief': some beliefs are doxastically rational, while others are not. The problem that the Revisionist faces is to explain how all-around irrational beliefs are any different from secondary cognitive attitudes. To explain this, I will need to make a slight detour to investigate these secondary cognitive attitudes.

2.1.1 Secondary Cognitive Attitudes

Secondary cognitive attitudes are imagination, assumption, acceptance, and the like (Velleman, 2000).[13] Secondary cognitive attitudes and belief can share contents: one can actually simultaneously believe a proposition and also imagine it (see Leslie, 1994), even

---

[13] For an account of these attitudes, see, *inter alia,* van Fraassen (1980), Stalnaker (1984), Bratman (1987), Rey (1988), Cohen (1992), and Van Leeuwen (2009; 2014).

though it has been convincingly argued that imagination is a self-standing cognitive attitude (cf. Peter Langland-Hassan, 2012). Furthermore, contrary to the claim that what motivates one to act is *always* a desire-belief combination, there are now convincing arguments claiming that supposition, imagination, etc., can access our inferential processes and theoretical reasoning in a belief-like fashion (e.g., Currie, 2002, Currie and Ravenscroft, 2002; Nanay, 2016; Van Leeuwen, 2009; 2014). Borrowing from David Velleman's well-known example, when a child pretends that he is an elephant, his imagining "I am an elephant" motivates the child to behave *as if* he were an elephant (Velleman, 2000; cf. Nichols and Stich 2000). This seems to indicate that secondary cognitive attitudes may also respond to evidence, at least in some restricted sense: within a game of make-believe, a child's imagining may respond to what counts as evidence *in that game* and may be updated accordingly. It is also plausible that we are prone to (imaginatively) subscribe to many of the consequence of the propositions we imagine (Berto 2017; Byrne 2016). This picture goes beyond imagining: we often also suppose, accept, or assume things to be thus-and-so, and act accordingly. If I accept a proposition that I do not believe for the sake of argument, I will be motivated to accept at least some of the propositions following from it.

Although they enjoy some belief-like features, it is widely recognized that secondary cognitive attitudes are immune to facts, compartmentalized, encapsulated, and highly context sensitive. Unsurprisingly, then, imagination has typically been described as the attitude that, unlike belief, takes little interest in the truth (Velleman, 2000). Furthermore, because secondary cognitive attitudes tend to operate in compartments, they admit contradictions (e.g., I can accept one thing in one context and the opposite in another), and they can also be "patchy" and poorly inferentially-integrated (Levy, 2015). Thus, while

imaginations *do* produce some inferences, imaginings tend only to produce new, inferentially-related imaginings within the imaginative context. Similarly, if I *accept* that my client is innocent, I may do so regardless or even in spite of the evidence I have. While my acceptance will have effects on the things I say and do at work, the accepted proposition will not normally spill over into my private life. For similar reasons, imagination and kindred cognitive attitudes will not motivate us across the board, namely beyond the contexts in which they were introduced (Van Leeuwen, 2009).[14]

2.1.2 Back to the Objection

Encapsulation, poor integration, and limited sensitivity to actual evidence are the marks of those cognitive attitudes that are not beliefs. What, then, about belief itself? The Strong Traditionalist has a straightforward story to tell us about what makes belief different from secondary attitudes. O'Brien (2005), for example, points out that beliefs are expected to range across the board, having an impact in different situations and contexts (see also Noordhof, 2001: 253). That is, a belief that p can enter reasoning whenever *p* is relevant qua premise, irrespective of the context. This also means that beliefs are largely consistent and coherent globally, whereas the effects of imagination and the rest are "short term" (Currie and Ravenscroft, 2002: 16).

---

[14] This position is also echoed in Bratman's (1990) theory of "acceptance in a context". On Bratman's view, acceptances, like beliefs, function as premises in practical reasoning and deliberation. However, the power of an acceptance is limited to the context in which it was introduced.

The Revisionist, of course, disagrees: she believes that belief is at times as fragmented and encapsulated as any secondary cognitive attitude. On her view, it is possible for a belief to have only limited effects. The worry I have, however, is that this threatens our ability to functionally distinguish mental attitudes, given that now belief loses its key functional markers and behaves similarly to imagination and kindred cognitive attitudes. As a result of this demarcation-problem, the Revisionists will also face an extension-problem, since they risk inflating the category of belief to include things that are not in fact beliefs. Whereas the Strong Traditionalist could end up with a very limited number of beliefs, the Revisionist faces a parallel worry: being too permissive in admitting highly doxastically-irrational attitudes to the rank of 'belief', she will find the category of belief to be too densely populated.

2.2. Assertion is Not Enough

The Revisionist may think she has an obvious reply to the objection above: sincere assertion marks out the difference between belief and the other cognitive attitudes. Even for the Traditionalist, sincere assertion is a type of action to which belief can give occasion: if I believe that phlogiston does not exist, then I am also disposed to assert that proposition in the appropriate circumstances. The Revisionist goes a step further, however: she takes assertion to be *fundamental* to belief. In Matthew Boyle's own words, "the very existence of [an agent's] belief that P is constituted by her persisting assent to P" (2009: 143).[15] Patchy, and behaviorally- or inferentially-fragmented, attitudes are still beliefs provided they are

---

[15] Kaplan claims that belief is not a disposition to act in a certain way but a disposition to make a sincere assertion (see Kaplan 1996: 109). See also Cherniak (1996) where he offers an "assent view" of belief.

expressed through sincere assertion. Assertion is enough of an indicator of belief regardless of whether the agent then follows up on what she says. This position is often paired with the idea that a 'feeling of truth' sets belief apart from other attitudes. For instance, Bayne and Pacherie claim that: "[a]rguably, to believe P is to think that P is (or is likely to be) true when one reflects on it, whereas no such *sense* of truth accompanies the state of merely imagining P" (2005: 168, my emphasis).[16]

Obviously, sincere-assertion and feelings of truth cannot be *necessary* conditions for being a belief. Non-human animals and babies have beliefs but they arguably lack the ability to assert or have any "sense of truth". In addition, unconscious beliefs do not give rise to assertions at all, and it is not clear they characteristically produce any epistemic feeling. This is not a concern of the Revisionist, however: she will point out that all she needs is to say is that assertion is a sufficient, but not a necessary, condition for being a belief. Although it is true that some beliefs will never produce assertion, if a proposition is sincerely asserted and felt to be true, then that is enough evidence that the agent has the corresponding belief.

Unfortunately, it is difficult to say what a sincere assertion *is* without making the – in this context, at least – unhelpful observation that an assertion is sincere if and only if the speaker actually believes its content (Searle, 1969: 65). Indeed, it should be obvious that uttering an indicative statement or sentence is not enough for that utterance to constitute an assertion. Actors on the stage and novelists in books say and write things that are not meant to be assertions even though they sound or look like assertions. In response, a popular view,

---

[16] Cohen (1992: 4) explains, "belief that p is a disposition, when one is attending to the issues raised, or items referred to, by the proposition p, normally to *feel it true that p* and false that not-p, whether or not one is willing to act, speak, or reason accordingly" [my emphasis].

prominently championed by Brandom (1994), claims that asserting is akin to undertaking a normative commitment. When asserting that, e.g., global warming is real, you are – as Brandom would put it – making a move in "the game of giving and asking for reasons", and you have committed yourself to supporting your claim. What's more, you are also committed to acting in a way that is consistent with your ideas about global warming. This means that the act of assertion is inextricability linked with the very notion of doxastic rationality: if you are overtly flouting doxastic, rational norms, then your utterance ceases to count as asserting. Overall, this view does not help the Revisionist because it indicates that normative standards set limits for what can count as an assertion. And this is precisely the claim the Revisionist wants to reject.

To avoid this dead-end, the Revisionist may simply insist that if a speaker is expressing what *feels true to her* then she is thereby sincerely asserting it. Thus, what sets belief apart from imagination is that when we believe that p, we will tend to assert it (when the circumstances are right, of course). However, I doubt that sincere assertion – even as an expression of this feeling of truth – reliably tracks belief. Until recently, it has virtually been a truism in philosophy that an assertion is sincere if and only if it expresses a belief (or some kind of feeling of truth). However, this view is losing ground since it seems that one can assert something that feels true while mistakenly taking herself to believe it (Peacocke 1998: 90; Schellenberg 2005). What this means is that an entrained proposition *p* may feel true to an agent S even though S does not really believe that p is true. This possibility is fully consistent with recent findings regarding people's failure to introspect: people are surprisingly prone to misattribute mental states to themselves (Carruthers 2011; Lawlor 2003; Scaife 2014). In a recent paper, Levy (2018) argues that cases in which a speaker

sincerely asserts that p but does not act on it tend to be cases in which she *mistakenly* believes to believe that p. The conscious feeling of truth and the sincere assertion she expresses indicate that *she takes herself to believe* it, not that she actually believes it (as also pointed out by Currie, 2000). Similarly, Richard Dub (2017) has argued that secondary cognitive attitudes (specifically 'acceptances') could at times be "generated by pathologically powerful cognitive feelings", and these powerful cognitive feelings can be sincerely expressed through assertions. The label 'cognitive feelings' refers to a broad category of emotion-like states including feelings of familiarity, feelings of knowledge, feelings of rightness, and that 'tip-of-the-tongue' feeling. Dub (2017: 50) makes a convincing case for the claim that, at least sometimes, these feelings:

> do not directly influence beliefs: they influence acceptances. A powerful
>
> feeling that p leads us to immediately accept p without our coming to
>
> believe that p.

When overtly expressed, these feeling-induced acceptances may well sound like beliefs. However, given the multiplicity of states that may be produced by cognitive feelings, "it is not introspectively clear to a speaker which mental state he or she is expressing at any one time" (Dub 2017: 45). Hence, this suggests that some non-doxastic attitudes produce sincere assertions and it is hard to set these cases apart from assertions that express beliefs. This casts a shadow over the Revisionist's claim that sincere assertion is the mark of belief: assertion simply seems insufficient in setting beliefs apart from other cognitive attitudes.

3. Moderate Traditionalism[17]

I will now present a position called Moderate Traditionalism. This view avoids both the Too Few Objection and the Too Many Objection by arguing that beliefs need to show a minimal degree of doxastic rationality:

Moderate Traditionalism: mental attitude A is a belief only if A is (at least) minimally rational.

I call this 'minimal' rationality for two reasons: it surfaces only when a certain level of detection (not necessarily conscious awareness) is in place; and it specifically concerns internal coherence. Accordingly, beliefs are minimally rational in the sense that they respond to *perceived* irrationality by reestablishing coherence.[18] I am not trying to say beliefs are in

---

[17] Possible defenders of this view are Davidson (1984), Dennett (1989), Pettit and Smith (1996) and Bayne and Pacherie (2005). Cherniak (1996) talks about 'minimal rationality' but he means something slightly different.

[18] The notion of coherence here loosely refers to what some philosophers call 'structural rationality'. Agents are said to be rational if their beliefs are coherent, namely if they conform to (formal and non-formal) standards of structural rationality (see Scanlon, 2007). Here are a couple of examples:

-Modus Ponens: Rationality requires of S that [S believes that q, if S believes that p, and S believes that if p then q, and S attends to the question whether q]

-Practical non-formal standard: Rationality requires of S that [S intends to do X, if S believes that she ought to do X]

general coherent. However, when incoherence is detected, standards of rationality are applied to re-establish it. Furthermore, clear failures to re-establish coherence are accompanied by a sense of distress. Note that this rational baseline can be guaranteed through an *automatic* process. Thus, complex cognitions are *not* necessary for preserving minimal rationality.

Against Strong Traditionalism, then, I do not expect rationality to be consistently apparent and manifested. Moreover, contrary to Revisionism, I expect a certain level of rationality in our beliefs: when belief fragmentation and irrationality are detected, there will be a contrary push for unity. I will also argue that no such reaction is expected in the case of secondary cognitive attitudes, which are governed by a decoupling mechanism that fosters fragmentation (Leslie, 1987).

Note that this view will be preferable over its competitors since it avoids the two objections mentioned above. Moderate Traditionalism avoids the Too Many Objection by clearly defining a way to set beliefs and the other cognitive attitudes apart without inflating the category 'belief'. What's more, Moderate Traditionalism is fully compatible with the evidence that some of our beliefs are irrational because it sets the bar for belief's rationality requirement *lower* than does Strong Traditionalism. This view is thus immune to the Too Few Objection as it is able to describe how our beliefs do in fact behave.

These standards are usually intended as having a wide-scope formulation. That means that, e.g., Modus Ponens can be satisfied either by coming to believe that q or by ceasing to believe that p (or that if p then q). Also, they need to avoid or resolve specific conflicts locally rather than globally (Kolodny, 2005: 516).

3.1. A Test Case

Before arguing that belief is minimally rational, it would be useful to focus on a case of a belief-like attitude that – on my view – should not be classified as a belief:

Superstition: Andrew seems persuaded that "if we raise our wine glasses in a toast, but we don't look into each-others' eyes as the glasses clink, we will be doomed to seven years of bad sex". Andrew has little evidence for this claim and has no idea what causal or explanatory relation there would be between bad sex and failing to look someone in the eyes. In fact, he tells us that he *knows* this superstition is in fact false (Huddleston, 2012: 209). Nevertheless, he hopes to convince us that this peculiar, superstitious mental state still counts as a belief, albeit a "naughty" one.

Andrew's avowed attitude is immune to counter-evidence and is mostly inferentially-isolated (e.g., even if we forget once to look into each-others' eyes as our glasses clink, I doubt we *really* expect to have bad sex for seven years). This superstition also fails to "give rise to the behaviors that we would normally expect to come in tow with certain beliefs" (Huddleston, 2012: 220). Andrew is himself acutely aware of these failings. A Revisionist such as Andrew Huddleston, though, argues that superstition in fact depicts a belief. Huddleston claims that the example above shows that the link between belief and rationality is flexible: the doxastically-irrational and recalcitrant attitude that characterizes superstition is just an extreme case of 'belief gone bad', and stands on a continuum with the other irrational beliefs that we can locate in the literature. Huddleston (2012: 216) furthermore explains that superstitious beliefs are "naughty beliefs", which "can sneak into the belief

game for free". All games have some unconventional moves, he claims: superstitious beliefs are the outliers in the belief-game.

Unfortunately, Huddleston's diagnosis of superstition runs into the problem of offering a convincing way to demarcate these naughty beliefs from secondary cognitive attitudes: given their characteristic inputs and outputs, naughty beliefs look awfully like assumptions and imaginations. What's more – and this is the key point here – Andrew does not seem to be particularly bothered by the oddity of his (alleged) belief: he feels little pressure to change things, even though he is aware that his attitude is blatantly irrational. In the remainder of the paper, I will present an argument for why Andrew's superstitious attitude is not a belief, and why we should expect beliefs to be at least minimally rational in a way that superstition is not.[19]

3.2 Belief's Minimal Rationality

Moderate Traditionalism – the view that beliefs are at least minimally rational – boils down to the following: to know whether an attitude is a belief, scrutinize how the attitude and the system react to *detected* irrationality and internal incoherence. If you see discomfort and coping strategies to avoid conflict, you are bearing witness to a belief. Thus, one of the defining features of belief is a certain *type* of irrationality-avoidance. By contrast, *an attitude is not a belief if it is impervious to changes even when the subject detects that her cognitive attitude breaches doxastic rationality*. Secondary cognitive attitudes are not subject to doxastic requirements since they are actually expected to be globally incoherent, evidence-

---

[19] Ichino (2019) also maintains that superstitions are not beliefs. And Currie and Jureidini (2002) use a similar argument to support the view that delusions are not beliefs.

unresponsive, and behaviorally inert. As a result, they behave differently, do not try to cope with doxastic irrationalities, and do not create the same emotional discomfort. That is, they are *not* minimally doxastically-rational attitudes.

Let us first look into belief's minimal rationality. The claim that belief is minimally rational seems to be intuitively right. Here is an example in Lewis (1982) in which an agent (aka himself) has the following three inconsistent beliefs: he believes that some street. runs North-South; he believes that the railroad tracks run East-West. He is also convinced that the street and the railroad tracks run in parallel. What is going on here, according to Lewis, is that the agent harbors contradictory beliefs because he fails to bring those beliefs to bear. However, the expectation is that the minute he recognizes what is going on, the agent will feel the immediate need to avoid this inconsistency. It would be *very* surprising to us if he were to do nothing and live happily with this odd set of beliefs. This is because being aware of harboring irrationality actually hurts. When that happens, we feel a strong sense of discomfort and are moved to change our attitudes in order to eliminate conflict.

So far, I believe this is all very intuitive. I will now refer to three possible psychological mechanisms to make sense of these intuitive ideas. What follows is *not* meant to be an exhaustive list, of course, and some of these mechanisms may seem to offer competing explanations. Nevertheless, each mechanism can be used to make sense of various beliefs' irrationalities and different irrationality-resolution strategies. And that is all I need here.

[ONE] A fragmentary mental architecture and belief storage system is thought to be among the possible reasons why beliefs can be encapsulated, inconsistent, and irrational. Contrary to the idea that beliefs are stored in an interconnected web-like structure, the

hypothesis that belief-tokens are stored in separate units explains why some of our beliefs are irrational in the ways I mentioned above.

There are various ways to explain fragmentation, but interestingly for my purposes here, at least one version of this fragmentation-view points out that, *within* the same fragment, inconsistencies are "automatically resolved", and "if two fragments that contain inconsistent information are coactive they will be rendered consistent" (Bendana and Mandelbaum, forthcoming).[20] This suggests that what allows irrationality to persist is in part the fact that fragments are often *not* simultaneously activated.

What's more, on this view inconsistency-resolution unfolds "along the lines predicted by dissonance theory". Let us see what that means. Research on dissonance indicates that we feel tension when we discover that our attitudes are in conflict, and so we employ various strategies to dissipate the tension. As Festinger explained, "[c]ognitive dissonance […] leads to activity oriented toward dissonance reduction just as hunger leads toward activity oriented toward hunger reduction" (see Harmon-Jones, 2000; Van Veen et al., 2009). On a popular model, viz., the action-based model of dissonance, the reason we feel discomfort in the face of conflicting cognitions is owing to their action-producing role: conflicting but simultaneously activated attitudes motivate divergent actions, undermining one's ability to engage in effective and resolute decision-making (Harmon-Jones, Amodio, & Harmon-Jones, 2009).[21]

---

[20] Bendana and Mandelbaum call this principle "synchronization".

[21] On a different view, dissonance's function is to preserve the self, since cognitive dissonance is especially experienced when incoming information conflicts with self-views (Aronson, 1968, 1969). Also, Mandelbaum (2019) maintains that dissonance functions as an "immune system" to preserve some of our core beliefs and keep them in place.

The cognitions involved in action-motivation are what philosophers call 'belief' and they produce discomfort when they do not line up with our actions. In fact, there is evidence that detected behavioral conflicts produce dissonance (Festinger and Carlsmith, 1959). How do we fix behavioral dissonance? Well, one way to eliminate conflicts between action and value is to change one's preference(s). For instance, results show that children think less of a toy if they are told they cannot play with it (Aronson and Carlsmith, 1963). Similarly, agents are likely to decrease their opinion of others when they are forced to show hostile behavior towards them (Davis and Jones, 1960; Glass, 1964). When you cannot change your actions, explains Festinger, you instead change how you see the world.[22]

A similar dissonance-avoidance mechanism presumably underlies the way in which we approach contrary evidence: elective evidence-seeking strongly suggests, as recently put by Mandelbaum (2019: 153), that, when holding fixed beliefs, "encountering disconfirming evidence hurts (and encountering confirmatory information feels good)". So, we tend to avoid this type of dissonance to avoid "psychological discomfort". Thus, a liberal might resist being exposed to conservative news outlets to avoid confronting evidence that may enter into conflict with what she believes (Sawicki et al., 2011; see Smith, Fabrigar, & Norris, 2008 for a review of how people tend to avoid dissonant information). This avoidance mechanism

---

[22] There is evidence that the phenomenon of dissonance is also present in children and non-human animals (Aronson and Carlsmith, 1963; Friedrich & Zentall, 2004; Egan, et al., 2007). This suggests that dissonance-detection and reduction may not always require metacognitive "high-level capacities" and second-order thoughts (Egan et al., 2007: 982). Consistently with this, Lieberman et al. (2001) suggest that the detection-reduction of cognitive dissonance is a fairly automatic process. Relatedly, Zanna et al. (1973) indicate that the discomfort associated with dissonance and ambivalent attitudes increases when the subject is consciously aware of the conflict.

seems related to the strength with which an attitude is held (Brannon, Tagler, and Eagly, 2007). Thus, another way to cope with perceived incoherence is to reinforce one's credence in the truth of some believed proposition, so as to fend off the cognitive impact of incoming counter-evidence (Lord, et al., 1979). There is at least some evidence of a "backfire effect" (Mandelbaum, 2018): subjects admit to having received counter-evidence for something they believe and this actually reinforces their certitude (Nyhan and Reifler, 2010; but cf. Tappin, & Gadsby, 2019; Wood, & Porter, 2018). Although these strategies may either produce or preserve beliefs, which are in fact epistemically irrational, they nonetheless manage to avoid internal conflicts between attitudes: apparent coherence is restored by upping one's credence in the truth of certain propositions while trivializing the disconfirming evidence (Festinger, 1957: 264; Simon, et al 1995).

[TWO] One need not posit a fragmented infrastructure of belief-storage to find evidence for belief's minimal rationality. On a widely-shared picture of how beliefs come to be formed and updated, the primary sources of beliefs' irrationality are those quick and dirty cognitive processes that are part of so-called "system 1". By contrast, the mechanism behind irrationality-reduction is analytic reasoning itself, or "system 2" (Kahneman and Tversky, 1979). On this view, people are irrational (indeed stubbornly so) largely on account of being cognitively "lazy", preferring to rely on fast-thinking processes (Pennycook & Rand, 2019b). This indolent attitude can be overcome by prompting agents to engage more in analytical reasoning. This is confirmed by mounting evidence that, in some cases, human beings are prone to revise their opinions and beliefs *once they attend* to the fact that their beliefs are unsupported or disproven. Indeed, studies seem to indicate that we can nudge subjects in overcoming their confirmation biases, for instance, by pushing them to consider evidence

against the views they hold (Schwind et al., 2012) and thus to engage in analytical reasoning (Pennycook & Rand, 2019a). Another study suggests that, in the presence of a low metacognitive "feeling of rightness", subjects are more prone to change their minds, or to at least engage in reflective analytical reasoning (Thompson, et al., 2011). This indicates that metacognitive engagement and attention may lead to the employment of reflective or analytical reasoning, as well as rational adjustments in one's belief system (Ranney & Clark, 2016).[23]

[THREE] Alongside evidence showing that slow, analytical reasoning can drive us to overcome deep-seated irrational tendencies, there is also evidence that a form of reasoning called "argumentative reasoning" pushes us to seek a certain degree of coherence and avoid perceived irrationalities. Mercier and Sperber (2011) have argued that argumentative reasoning has evolved with a social-communicative function: forming arguments advances our social status by convincing others of the correctness of our attitudes. Thus, on their view, reasoning has a *post hoc* rationalizing function in creating seemingly coherent arguments for attitudes we already hold.[24] By contrast, since we cannot defend patently incoherent or irrational attitudes, these are the attitudes that put us at risk of losing social standing. This presumably creates some pressure on us to come up with ways of re-establishing our sense

---

[23] Bortolotti (2010: 87) contends that recent evidence in psychology shows that beliefs do not often get revised even when their rational failure becomes apparent (Tversky and Kahneman, 1983). It is unclear, though, whether the subjects in those studies are willing to accept that they made a mistake or that their beliefs are irrational. Also see, e.g., Ranney & Clark (2016) for different results.

[24] Similarly, Campbell & Kumar (2012) argue that people often engage in "consistency reasoning", intended here as a separate cognitive mechanism, which does the job of coordinating slow, analytical thinking with fast, emotion-based cognitive processes.

of coherence. The discomfort we feel upon confronting our irrationalities could be the result of this pressure.

It is important to clarify that, once again, irrationality-resolution strategies may dissipate a *perceived* sense of dissonance rather than actually making us rational. Indeed, attempts at re-establishing coherence may be the source of *more* overall irrationality when they are the result of motivated reasoning. However, I contend that any pressure on internal coherence is itself an expression of *rational constraints* operating on beliefs in a way that they do not operate on other cognitive attitudes.

To support this view, there is now substantial evidence that at times we do indeed reason for argumentative and social purposes (for a review of the evidence see Mercier and Sperber, 2011). This kind of motivated reasoning could also be the driving force behind belief polarization (Kahan, 2016[25]). Similarly, Haidt (2001) has argued that moral reasoning is mostly responsible for concocting justifications to convince others of the goodness of our deep-seated moral intuitions. And motivated reasoning could perhaps also be seen as the driving-force behind cases of attitude-blindness, whereby people "forget" their original decisions or choices and change their attitudes to square them with the reasons currently available to them (Hall et al., 2012). Finally, confabulation – the creation of false and unsupported explanations – is used to resolve the conflict between attitudes, whereby new, confabulated attitudes are created to square one's actions with one's beliefs in order to avoid conflicting cognitions. In fact, confabulation often occurs when subjects are asked to explain

---

[25] Kahan (2016) maintains that motivated reasoning is responsible for altering the strengths of incoming evidence when this is seen as in tension with the convictions of our social or political group.

and/or justify behavior or attitudes that are – unbeknownst to them – the result of unconscious processes (Nisbett & Wilson, 1977; Gantman et al., 2017; Greene, 2014). This suggests that confabulation is yet another coping mechanism through which we try to avoid the psychological and social burden of facing our irrationality (Carruthers, 2011: 356-365).

3.3 Decoupling without Irrationality-Avoidance

In the previous section, I reviewed the evidence suggesting that we have a number of psychological mechanisms that react to our perceived irrationality, and which try to dissipate the tension thereby created in us. As explained, irrationality-avoidance and resolution are often associated with the discomfort we feel when our cognitions are perceived to be fragmented, isolated, inconsistent, and so on. But not all cognitions are created equal. Indeed, we typically do not find this reaction in cases of imagination or acceptance. Rather, when it comes to imagination, it seems we readily tolerate contradictions, dissonant behavior, and evidence-insensitivity. This is not to say that secondary cognitive attitudes are free from constraints. But whatever rational requirements might oversee them, they do not react *in the same way as belief does.* Contradictory imaginations, for instance, are tolerated, and are not perceived 'as a failing', provided that they belong to different "imaginative episodes". As Bayne puts it (2010: 333):

> [t]he norm of consistency applies to belief in a global fashion—to have
>
> inconsistent beliefs is to exhibit some kind of failing qua believer—but it
>
> applies to imagination only within particular imaginative episodes.

When applied to imaginative episodes, fragmentation is not only completely acceptable but is an *asset* allowing us to engage in complicated reasoning by keeping track of multiple possibilities at the same time. In fact, these secondary attitudes are subject to a *decoupling mechanism* allowing us to form imagined narratives that are cognitively distinct (that is, they occupy a distinct cognitive space) from things that are perceived, believed, or thought to be true, or from other sets of narratives (Leslie, 1987; Perner, 1991). Although coherence is usually maintained within imaginative episodes, it is always possible to consider multiple narratives at the same time, thus permitting incoherent attitudes to co-exist. This is what makes this type of fragmentation different from the fragmentation that plagues our beliefs (see Currie and Ravenscroft, 2002: 16, 176). That is, while a decoupling mechanism may be at work in imagination and kindred cognitive attitudes, this mechanism does *not* apply to belief. Even if beliefs can stand in contradiction, and are at times compartmentalized, this is a state that the *cognitive system perceives to be anomalous*, and will try to remedy upon recognition of the problem. Although fragmentation is possible, belief's push for coherence applies indiscriminately and does not admit compartmentalization. This is what renders belief rational, at least in a minimal sense.[26]

REFERENCES

Adler, J. E. (2002). *Belief's own ethics.* Cambridge, MA: MIT Press.

Ajzen, I., & Fishbein, M. (2005). The Influence of Attitudes on Behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-221). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Anderson, C. A. (1995). Implicit personality theories and empirical data: Biased assimilation, belief perseverance and change, and covariation detection sensitivity. *Social Cognition,* 13, 25-48.

Alfnes, F., Yue, C., and Jensen, H. H. (2010). Cognitive dissonance as a means of reducing hypothetical bias. *Eur. Rev. Agric. Econ.* 37, 147–163.

Aronson, E. (1968). Dissonance theory: Progress and problems. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcombe, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 5-27). Skokie, IL: Rand-McNally.

Aronson, E. (1969). The theory of cognitive dissonance: A current perspective. *Advances in Experimental Social Psychology, 4,* 1-34.

Aronson, E., and Carlsmith, J. M. (1963). Effect of the severity of threat on the devaluation of forbidden behavior. *J. Abnorm. Soc. Psychol.* 66, 584–588.

Bendana J. and E. Mandelbaum (forthcoming). The Fragmentation of Belief. in *The Fragmentation of Mind* (eds. Dirk Kindermann, Cristina Borgoni, & Andrea Onofri) Oxford: Oxford University Press.

Bayne, T. (2010). Delusions as Doxastic States: Contexts, Compartments, and Commitments. *Philosophy, Psychiatry, and Psychology* 17: 329–336.

Bayne T., and Hattiangadi A. (2013) Belief and Its Bedfellows. In: Nottelmann N. (eds) New

Essays on Belief. Palgrave Macmillan, London

Bayne, T., and E. Pacherie (2005). In defence of the doxastic account of delusions. *Mind & Language* 2:163–88.

Berto, F. (2017) Impossible Worlds and the Logic of Imagination, *Erkenntnis*, 82: 1277–1297

Boyle, M. (2011). 'Making Up Your Mind' And The Activity Of Reason. *Philosophers' Imprint* 11: 1-24.

Bortolotti, L. (2010). *Delusions and Other Irrational Beliefs*, Oxford: Oxford University Press.

---- (2012). In Defense of Modest Doxasticism about Delusions, *Neuroethics* 5: 39–53.

Brandom, B. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press, Cambridge.

Broome, J. (2013). Rationality through reasoning. Oxford: Wiley Blackwell.

Bratman, M. E. (1992). Practical reasoning and acceptance in a context. *Mind*, 101:1–15.

Campbell, J. (2001). Rationality, meaning, and the analysis of delusion. *Philosophy, Psychiatry, & Psychology*. 8: 89–100.

Campbell, R., & Kumar, V. (2012). Moral Reasoning on the Ground. *Ethics,* 122: 273-312.

Brannon, L. A., Tagler, M. J., & Eagly, A. E. (2007). The moderating role of attitude strength on selective exposure to information. *Journal of Experimental Social Psychology*, 43: 611-617.

Broome, J. (2007). 'Wide or Narrow Scope?,' *Mind*, 116/462: 359-370.

 ---- (2013). *Rationality Through Reasoning*. Chichester: Wiley-Blackwell.

Byrne, R.M.J. (2016). Counterfactual thinking. In A. Kind (ed). *The Routledge Handbook of Philosophy of Imagination*. London: Routledge. Pp 339-352.

Carruthers P (2011). *The opacity of mind: An integrative theory of self-knowledge*. Oxford: Oxford University Press.

Cherniak, C. (1986). *Minimal Rationality.* MIT Press.

Clore, G. and Gasper, K. (2000). Feeling is believing: Some affective influences on belief. In Frijda, N., Manstead, A., and Bem, S. (eds). *Emotions and Beliefs: How Feelings Influence Thoughts.*

Cohen, L. J. (1992). *An Essay on Belief and Acceptance.* New York: Clarendon Press.

Currie, G. (2000), Imagination, Delusion and Hallucinations. *Mind & Language*, 15: 168-183.

---- (2002). Imagination as Motivation. *Proceedings of the Aristotelian Society*, 102:201-16.

Currie, G. and Jureidini, J. (2002): Delusion, rationality, empathy. Philosophy, Psychiatry & Psychology, 8/2, 159–62.

Currie, G & Ravenscroft, I. (2002). *Recreative Minds: Imagination in Philosophy and Psychology.* Oxford University Press.

Davidson, D. (1963). Actions, Reasons, and Causes. *The Journal of Philosophy,* 60, 685-700.

---- (1980). *Essays on Actions and Events.* Oxford: Oxford University Press.

---- (1984). *Inquiries Into Truth And Interpretation.* Oxford University Press.

Davis, K and Edward J. (1960). Changes in Interpersonal Perception as a Means of Reducing Cognitive Dissonance. *Journal of Abnormal and Social Psychology* 61 (3): 402–10

de Sousa, R. (1971). How to give a piece of your mind: or, the logic of belief and assent. *Philosophical Review*, 35: 52–79.

Dennett, D. (1987). *The Intentional Stance.* MIT Press.

DiLalla, L.F. and Watson, M.W. (1988). Differentiation of Fantasy and Reality: Preschoolers' Reactions to Interruptions in Their Play. *Developmental Psychology* 24: 286–291.

Dub, R. (2017), Delusions, Acceptances, and Cognitive Feelings. *Philos Phenomenol Res*, 94: 27-60.

Egan, A. (2008). Seeing and Believing: Perception, Belief Formation, and the Divided Mind. *Philosophical Studies* 140: 47-63.

Egan, A. & Elga, A. (2005). I Can't Believe I'm Stupid. *Philosophical Perspectives,* 19: 77- 94.

Egan, L. C., Santos, L. R., and Bloom, P. (2007). The origins of cognitive dissonance: evidence from children and monkeys. *Psychol. Sci.* 18, 978–983.

Fabrigar, L. R., Petty, R. E., Smith, S. M., & Crites, S. L., Jr. (2006). Understanding knowledge effects on attitude-behavior consistency: The role of relevance, complexity, and amount of knowledge. *Journal of Personality and Social Psychology,* 90: 556-577.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, *58*, 203-211.

Fodor, J. (1985). Fodor's guide to mental representation. *Mind*, 94, 76–100.

---- (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT/Bradford.

---- (1990). *A Theory of Content and Other Essays*. Cambridge: MIT/Bradford Press.

Friedrich, A.M., & Zentall, T.R. (2004). Pigeons shift their preference toward locations of food that take more effort to obtain. *Behavioral Processes*, 67: 405–415.

Frankish, K. (2007). Partial belief and flat-out belief. In F. Huber, & C. Schmidt Petri(Eds.), *Degrees of belief: An anthology*. Oxford: Oxford University Press.

---- (2004*). Mind and Supermind*. Cambridge: Cambridge University Press

---- (2011). Delusions, Levels of belief, and non- doxastic acceptances. Neuroethics doi:10.1007/s12152- 011-9123-7.

Frankfurt, H., (1992). The Faintest Passion. *Proceedings and Addresses of the American Philosophical Association*, 66, 3. 5-16.

Gantman, A. P., Adriaanse, M. A., Gollwitzer, P. M., & Oettingen, G. (2017). Why did I do that? Explaining actions activated outside of awareness. *Psychonomic Bulletin & Review* 24: 1563-1572

Gendler, T. S. (2000). The Puzzle of Imaginative Resistance. *The Journal of Philosophy*, XCVII:55-81.

---- (2003). On the relation between pretense and belief. In D. M. Lopes & M. Kieran (Eds.), *Imagination, philosophy and the arts.* London: Routledge. pp. 125–141.

---- (2006). Imaginative Contagion. *Metaphilosophy*, 37: 1-21.

---- (2007). 2007. Self-Deception as Pretense. *Philosophical Perspectives*, 21: 231-258.

---- (2008). Alief and Belief. *Journal of Philosophy*, 105: 634-663.

Glass, David C. 1964. "Changes in Liking as a Means of Reducing Cognitive Discrepancies between Self-Esteem and Aggression." Journal of Personality 32 (4): 531–49.

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* 3: 137-162.

Golomb, C and Kuersten, R. (1996). On the Transition from Pretense Play to Reality, *British Journal of Developmental Psychology* 14: 203–217.

Greco, D. (2015). Iteration and Fragmentation. *Philosophy and Phenomenological Research.,* 91: 656–673.

Greene JD (2014) *Moral tribes: emotion, reason and the gap between us and them*. Atlantic Books, London

Guenther, C. L., & Alicke, M. D. (2008). Self-enhancement and belief perseverance. *Journal of Experimental Social Psychology*, *44*(3), 706-712. https://doi.org/10.1016/j.jesp.2007.04.010

Haidt J., (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108: 814-825

Hall L, Johansson P, Strandberg T (2012) Lifting the veil of moral- ity: choice blindness and attitude reversals on a self-transforming survey. *PLoS ONE* 7:e45457

Harmon-Jones, E., Amodio, D. M., & Harmon-Jones, C. (2009). Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict. In M. P. Zanna (Ed.), *Advances in experimental social psychology. Advances in experimental social psychology,* pp. 119-166. San Diego, CA, US: Elsevier Academic Press.

Helton, G. (2019), If You Can't Change What You Believe, You Don't Believe It. *Noûs*.

Hirstein, W. and V. S. Ramachandran. (1997). Capgras Syndrome: A Novel Probe for Understanding the Neural Representation of the Identity and Familiarity of Persons. *Proceedings of the Royal Society of London* B264: 437–444.

Huddleston, A. (2012). Naughty Beliefs, *Philosophical Studies* 160: 209-222.

Hunter, D. (2011). Alienated Belief. *Dialectica* 65: 221-240.

Ichino A. (2018). Superstitious Confabulations. *Topoi* 1-15.

Kahan, D. M. (2016). The Politically Motivated Reasoning Paradigm, Part 1: What Politically Motivated Reasoning Is and How to Measure It. In *Emerging Trends in the Social and Behavioral Sciences* (eds R. A. Scott and S. M. Kosslyn).

Tversky, A., & Kahneman, D. (1982). Judgment under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and*

Biases (pp. 3-20). Cambridge: Cambridge University Press.

Kaplan, M. (1996). *Decision Theory as Philosophy*. Cambridge: Cambridge University Press.

Kolodny, N. (2005). Why Be Rational? *Mind* 114: 509–563.

---- (2008). The Myth of Practical Consistency. *European Journal of Philosophy,*16: 366–402.

Kunda, Z. (1990) The case for motivated reasoning. *Psychological Bulletin*. 108:480 – 98.

Lawlor K (2003) Elusive reasons: a problem for first-person authority. *Philos Psychol* 16:549–564

Langland-Hassan, P. (2012). Pretense, imagination, and belief: the Single Attitude theory. *Philosophical Studies* 159:155-179

Leslie, A. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50, 211–238.

---- (1987). Pretense and Representation: The Origins of "Theory of Mind". Psychological Review. 94. 412-426.

Levy, N. (2015). Neither Fish nor Fowl: Implicit Attitudes as Patchy Endorsements. *Nous*. 49: 800–823

---- (2018). You meta believe it. *Eur J Philos*. 26: 814– 826.

Lewis, D. (1982). Logic for Equivocators. *Nous* 16, 431-441. Reprinted in D. Lewis, *Papers in Philosophical Logic*, Oxford, Oxford University Press, pp. 97-110.

Lieberman, M.D., Ochsner, K.N., Gilbert, D.T., & Schacter, D.L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*, 12: 135–140.

Lord, et al. (1979) Biased Assimilation and Attitude Polarization: The Effects of Prior

Theories on Subsequently Considered Evidence, *Journal of Personality and Social Psychology*, 37, 11: 2098-2109,

Mandelbaum, E. (2014) Thinking is Believing, Inquiry: An Interdisciplinary Journal of Philosophy, 57:1, 55-96

--- (2016). Attitude, inference, association: On the propositional structure of implicit bias. *Noûs* 50: 629–658.

--- (2019). Troubles with Bayesianism: An introduction to the psychological immune system. *Mind Lang.* 2019; 34: 141– 157

Mercier, H., & Sperber, D. (2011). Why Do Humans Reason? Arguments for an Argumentative Theory. *Behavioral and Brain Sciences* 34: 57-74.

Meylan, A. (2014). Epistemic emotions: a natural kind? *Philosophical Inquiries* 2: 173–190.

Moran, R. (2001). *Authority and estrangement: An essay on self-knowledge*. Woodstock: Princeton University Press.

Nanay, B. (2016). The Role of Imagination in Decision-Making. Mind and Language, 31: 127–143.

Nichols, S. & Stich, S. (2000): A cognitive theory of pretense. *Cognition*, 74:115–147.

---- (2003). *Mindreading*. Oxford: Oxford University Press.

Newby-Clark, I. R., McGregor, I., & Zanna, M. P. (2002). Thinking and caring about cognitive inconsistency: When and for whom does attitudinal ambivalence feel uncomfortable? *Journal of Personality and Social Psychology*, *82*, 157-166

Nichols, S., Stich, S., Leslie, A. and Klein, D. (1996). Varieties of off-line simulation. In *Theories of Theories of Mind*. P. Carruthers and P. Smith (eds). Cambridge: Cambridge University Press.

Nisbett RE, Wilson TD (1977). Telling more than we can know: verbal reports on mental

processes. *Psychol Rev* 84:231–259

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior,* 32: 303–330.

Noordhof, P. (2001). Believe What You Want. *Proceedings of the Aristotelian Society* 101:247-265.

O'Brien, L. (2005). Imagination and the Motivational View of Belief. *Analysis*, 65: 55– 62.

Peacocke, C. (2000). Conscious Attitudes, Attention and Self-Knowledge, in *Knowing Our Own Minds: Essays on Self-Knowledge*, ed. C. McDonald, B. Smith and C. Wright (Oxford University Press).

Pennycook, G. & Rand, D. G. (2019) Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences.* 116 (7) 2521-2526

---- (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition.* 188, 39-50

Pettit P and Smith, M. (1996). Freedom in Belief and Desire. *The Journal of Philosophy*, 93: 429-449

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Ranney, MA, & Clark, D. (2016). Climate Change Conceptual Change: Scientific Information Can Transform Attitudes. *Topics in Cognitive Science*, 8(1), 49-75.

Rey, G. (1988). Toward a computational account of akrasia and self- deception. In McLaughlin, B. P. and Rorty, A. O., editors, *Perspectives on Self-Deception*. University of California Press.

Reimer, M. (2010). Only a philosopher or a madman: Impractical delusions in philosophy and

psychiatry. *Philosophy, Psychiatry, & Psychology* 4:315–28.

Rose, D., W. Buckwalter, and J. Turri. (2014). When Words Speak Louder than Actions: Delusion, Belief and the Power of Assertion. *Australasian Journal of Philosophy.* 92: 1–18.

Rose, D. et al. (2017). Behavioral Circumscription and the Folk Psychology of Belief: A Study in Ethno-Mentalizing. *Thought.* 6: 193–203.

Sawicki, V. *et al.* (2011) 'Seeking Confirmation in Times of Doubt: Selective Exposure and the Motivational Strength of Weak Attitudes', *Social Psychological and Personality Science*, 2: 540–546.

Scanlon, T.M. (2007). Structural Irrationality. In *Common Minds. Themes from the Philosophy of Philip Pettit*, edited by Geoffrey Brennan, Robert Goodin, Frank Jackson, and Michael Smith, 84–103. Oxford: Oxford University Press.

Schellenberg, S. (2013). Belief and Desire in Imagination and Immersion, *The Journal of Philosophy* 110: 497–517.

Scaife R (2014) A problem for self-knowledge: the implications of taking confabulation seriously. *Acta Analytica* 29:469–485

Schellenberg, J. L., (2005), *Prolegomena to a Philosophy of Religion*, Ithaca, NY: Cornell University Press.

Schwind, et al (2012). *Preference-inconsistent recommendations: An effective approach for reducing confirmation bias and stimulating divergent thinking*? Computers & Education. 58: 787-796,

Schwitzgebel, E. (2001). In-between believing. *The Philosophical Quarterly*, 5:76–82.

---- (2002). A phenomenal dispositional account of belief. *Nous*, 36: 249–275.

---- (2010). Acting contrary to our professed beliefs, or the gulf between dispositional belief

and occurrent judgment. *Pacific Philosophical Quarterly*, 91: 531–553.

Searle, J. (1969) *Speech acts*. Cambridge University Press.

Shah, N. and Velleman, J. D. (2005). Doxastic Deliberation , *Philosophical Review*, 114: 497-534.

Simon, L., Greenberg, J., & Brehm, J. (1995). *Trivialization: the forgotten mode of dissonance reduction*. *Journal of Personality and Social Psychology, 68*(2), 247-260.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Smith, S. M., Fabrigar, L. R., & Norris, M. E. (2008). Reflecting on six decades of selective exposure research: Progress, challenges, and opportunities. *Social & Personality Psychology Compass*, 2: 464-493.

Simon, H. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129-138

Smithies, D. (2012). The normative role of knowledge. *Noûs* 46: 265-288.

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Blackwell.

Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT.

---- (2002). Common Ground. *Linguistics and Philosophy* 25: 701-721.

Stein, E. (1996). *Without Good Reason*. Oxford: Clarendon Press.

Stephens, G. L. and Graham, G. (2004). Reconceiving delusion. *International Review of Psychiatry*, 16: 236–241.

Stich, S. (1978). Beliefs and subdoxastic states. *Philosophy of Science* 45: 499–518.

---- (1983). *From Folk Psychology to Cognitive Science*. Cambridge, Mass.: MIT Press.

---- (1990). *The Fragmentation of Reason*. Cambridge, MA: MIT Press.

---- (1991) Do true believers exist? *Proceedings of the Aristotelian Society*, Suppl. 65: 229–44.

Tappin, M & Gadsby, S. (2019). Biased belief in the Bayesian brain: A deeper look at the evidence, *Consciousness and Cognition*, 68:107-114.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L. & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331:1279–1285

Thompson, T., Prowse Turner , J. Pennycook G. (2011) Intuition, reason, and metacognition *Cognitive Psychology* 63 (3):107-140

Tversky, A. & Kahneman, D. (1983) Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90:293– 315.

Van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.

Van Leuween, N. (2009). The motivational role of belief. *Philosophical Papers*, 38: 219–246

---- (2014). Religious Credence is not Factual Belief. *Cognition,* 133: 698-715.

---- (2018). The Factual Belief Fallacy. *Contemporary Pragmatism* (eds. T. Coleman & J. Jong): 319-343.

Van Veen, V.; Krug, M.K.; Schooler, J.W.; Carter, C.S. (2009). Neural Activity Predicts Attitude Change in Cognitive Dissonance. *Nature Neuroscience*. 12:1469-74

Velleman, D. (2000). *The possibility of practical reason*. Oxford: Oxford University Press.

Viedge, N. (2016). Defending evidence-resistant beliefs. *Pacific Philosophical Quarterly* 99: 517– 537

Weinberg, J., & Meskin, A. (2006). "Puzzling Over the Imagination: Philosophical Problems, Architectural Solutions." In S. Nichols (Ed.), *The Architecture of Imagination*. Oxford: Oxford University Press.

Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt

behavioral responses to attitude objects. *Journal of Social Issues,* 25:41-78.

Williams, D. (2018). Hierarchical Bayesian models of delusion. *Consciousness and Cognition, 61*, 129–147.

Wood, T., & Porter, E. (2018). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior,*1:1–29.

Zanna, M. P., Lepper, M. R., & Abelson, R. P. (1973). Attentional mechanisms in children's devaluation of a forbidden activity in a forced- compliance situation. *Journal of Personality and Social Psychology, 28,* 355–359.

Zimmerman, A. (2007). The nature of belief. *Journal of Consciousness Studies*, 14: 61–82.