A Minimalist Threshold for Epistemically Irrational Beliefs

M.B. Ganapini

*** PENULTIMATE DRAFT – DO NOT CITE***
To appear in "What is Belief?"
Eric Schwitzgebel & Jonathan Jong (eds.)
Oxford University Press (forthcoming)

This paper aims to shed light on the nature of belief and provide support to the view that I call 'Minimalism'. It shows that Minimalism is better equipped than the traditional approach to separating belief from imagination and addressing cases of belief's evidence-resistance. The key claim of the paper is that no matter how epistemically irrational humans' beliefs are, they always retain a minimal level of rationality.

1. Two Traditionalist Views

Any analysis of the psychological nature of belief should provide cognitive science with tools to make predictions and offer compelling explanations concerning people's actions. In this paper, I do so by offering additional support to a view I here refer to as 'Minimalism'[1]. The main goal of Minimalism is to offer a way to distinguish belief from other similar attitudes such as imagination, acceptance and supposition. Furthermore, Minimalism promises to shed light on the nature of delusions and other irrational attitudes while setting clear expectations for when we should call an attitude 'belief'.[2] The focus here is not to vindicate

[1] I defended the full-fledged version of this view in Bergamaschi Ganapini (2020).

[2] Not everyone agrees on the feasibility of this kind of project. Schwitzgebel (2001) and Schellenberg (2013) argue that 'belief' is a vague term that stands on a continuum, and it is thus hopeless to try to clearly demarcate belief from other neighboring attitudes.

folk-psychology or to account for what people call 'belief'. In contrast, the explicit goal of my analysis is to provide useful tools for a mature science of belief.[3]

The reason why we need to adopt Minimalism is that the traditional approach has been struggling to make sense of important recent data coming out of psychology and cognitive science. It is standard in philosophy to see belief as the cognitive attitude that (i) is sensitive to the relevant evidence, (ii) is inferentially integrated with other beliefs and intentional attitudes, and it (iii) guides actions (when coupled with the appropriate conative attitude, in the right circumstances). In particular, when it comes to beliefs' evidence-responsiveness, the Traditionalist account points to the following as a necessary condition for being a belief:

(T-1): a subject S' cognitive state D is a belief that p only if D tends to respond to the evidence bearing on p

This picture has been widely endorsed in philosophy and boils down to the idea that belief is the kind of attitude that will be typically caused by some epistemic source (e.g. reasoning, perception) and will change according to the evidence one acquires (Currie and Ravenscroft 2002, Tenenbaum et al. 2011, Velleman 2000:277). That means that by and large beliefs are expected to be epistemically rational: they are produced in a rational way and/or they align with the evidence once formed. One of the merits of this approach is that it offers clear guidelines on how to distinguish belief from bordering attitudes such as imagination and acceptance (Van Leeuwen, 2009). Though they may cause action in some contexts, imagination and the like are usually completely evidence-irresponsive. They are not brought about by any source of knowledge and they steadily run contrary to the available evidence. This view is also often invoked to argue that delusions cannot be beliefs (Currie and Ravenscroft 2002).

Unfortunately, the problem with this old-fashion Traditionalist picture of belief is that it fails to account for the mounting evidence that a sizable chunk of our ordinary beliefs is epistemically irrational: some of our beliefs do not seem to respond to the available evidence

[3] As argued at length in Bergamaschi Ganapini (2020), I do not believe that sincere assertion is a sufficient indicator of belief, and I am skeptical that we should use people's introspection as a guide to track beliefs.

(Bortolotti 2010; Mandelbaum and Porot forthcoming). This raises a pressing worry for Traditionalism: if we adopt such a demanding view of belief, we are forced to apply the notion of belief only to a small group of (very rational) attitudes (Helton, 2018; Bergamaschi Ganapini 2020; see also Viedge 2018).[4]

These concerns have brought some to endorse a new version of Traditionalism which refers to the capacities (or abilities) subjects have to revise their beliefs. On this view, the focus is not on belief's actual behavior but on what belief is able to do *in principle*: even when behaving irrationally, beliefs are able to be rational. That is, proponents of this modified version of Traditionalism focus on belief's evidence-responsiveness in particular, and argue that belief is that one attitude for which a subject necessarily requires the ability or capacity to rationally respond to the evidence (Flores, 2021; Helton, 2018). If beliefs show evidence-insensitivity, it is because some "masking conditions" do not allow subjects to respond according to their abilities and capacities. Thus, this revised Traditionalism subscribes to the following necessary condition for being a belief:

(T-2): a subject S' cognitive state D is a belief that p only if S has the ability (or capacity) to revise D according to the evidence bearing on p

To be sure belief's evidence-responsiveness cannot be a matter of pure luck: it needs to be the result of the right type of mechanism or skills causing belief to align with the relevant evidence. In this way, (T-2) can make sense of the fact that some of our beliefs do not seem rational and do not respond to the evidence in the right way. When that happens, it is because of some condition that prevents the subject S from exercising her abilities.

[4] In addition, there is some evidence that we believe whatever proposition we entertain first irrespective of the evidence. The Cartesian model of our mental architecture comprises the widespread view that one can entertain a proposition, in imagination or perception, without first believing it. Believing (or disbelieving) is thus an active process that requires some effort and happens after one has entrained a proposition without believing it. In opposition to this standard model, Eric Mandelbaum has recently proposed the Spinozian model. As he explains it, "[p]eople do not have the ability to contemplate propositions that arise in the mind, whether through perception or imagination, before believing them. Because of our mental architecture, it is (nomologically) impossible for one to not immediately believe propositions that one tokens." (Mandelbaum, 2014: 61) The idea is that any proposition is automatically and passively believed first. Only after that initial assent, the belief may be revised.

Helton (2018) puts this in terms of psychological skills:

> "subjects are able, given their current psychological mechanisms and skills, to revise their beliefs in accordance with the evidence. 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." [5]

In a recent paper, Flores (2021) uses a similar idea to argue for why we should take (most) delusions to be as evidence responsive as ordinary beliefs are. On this view, evidence responsiveness is the capacity to respond to evidence by changing one's attitudes in an epistemically responsible way. When it comes to delusions, Flores argues that (some) subjects with delusions have the capacity to respond to the evidence. To explain the notion of 'capacity' Flores (2021: 6306) writes:

> "having the capacity to F involves successfully F-ing in specific conditions that suit that capacity. For example, what matters to whether one has the capacity to run a 40-min 10k is whether one does so when exerting serious effort, not injured, well-rested, highly motivated, and so on—even if one would fail to do so if a single one of these conditions is not met."

This account is surely a departure from the way the distinction between competence and performance is usually drawn. Problems with performance are usually linked to *momentary* mistakes. In contrast, *systemic* violations of rationality are considered to be a problem of competency and capacity (Stein, 1997:8). Contrary to this, we saw above that Flores' account of capacity explicitly allows for systematic deviations from norms of rationality. This allows Flores to open the door to the possibility that delusions have rational capacities even if they

[5] Helton (2018) maintains that without rational skills it is unclear how epistemic norms of revision can apply to believing. If we do not have the ability or a functioning mechanism to make our beliefs rational, we cannot be required to do so, for ought-implies-can reasons. However, this could actually be a reason for favoring the idea that there is no such a norm of belief or that that norm is evaluative (McHugh 2012, 2014; Bergamaschi Ganapini 2020b; Glüer & Wikforss, 2013).

are mostly relentlessly irrational. Flores points to a number of reasons to think delusions retain the capacity to rationally respond to the evidence, though that capacity is thwarted by some strong masking factors (e.g. strong somatic experiences, motivational factors). On this view, the presence of these factors explains why delusions do not rationally adjust to the evidence. In a similar fashion, some run-of-the-mill beliefs fail to respond to the evidence because conditions are not right for them to do so. That is, even ordinary beliefs may fail to align with the evidence if motivation, hunger, fatigue or other factors prevent subjects from exercising their capacities. In sum, this view offers a way to account for the evidence that belief is at times irrational while also defending the idea that humans have the capacity to be rational.

2. Where T-2 fails

Despite its appeal, I find T-2 quite unsatisfactory. For starters, it is worth pointing out that although (T-2) may initially seem in line with the spirit of (T-1), it in fact constitutes a substantial departure from the kind of rationalism put forward by the old-fashioned Traditionalism. Traditionalism, as the view of belief most championed in philosophy of mind, expects beliefs to show rationality (barring some extenuating circumstances). In contrast, (T-2) is compatible with beliefs behaving completely and openly irrationally (as long as this behavior can be explained by some masking condition).

This raises two sorts of worries. One is that some psychological masks may actually erode rationality, threatening the plausibility of (T-2). The second is that (T-2) leaves it undetermined what counts as a masking condition for belief vs. a condition that allows imagination to emerge.

Let's start from the top. A typical example of a mask is a piece of Styrofoam wrapped around a China pot: the latter would not break if dropped, even if the China is fragile (Bird 1998; Fara 2008). This is a mask that prevents an underlying ability to emerge: the ability is ready to yield its effects but an intervening factor prevents that (like an antidote, Bird 1998).

However, other types of masks intervene on the source of the very ability or capacity in question, as recognized in the literature (these are called also 'fink', Martin 1994). This opens

up the door for the first worry I mentioned above as a relevant chunk of psychological masks that cause irrationality, is of this latter kind: they do not simply stop rational skills from causing the right behavior, they actually infiltrate or remove the mechanisms responsible for our ability to respond to the evidence. Let's see one example:

Depressed Robert: Bob is so depressed that he completely fails to see how loved and lucky he is. Depression cripples his ability to respond to the evidence on anything related to his self-esteem. His depression is not just an impediment: it erodes his mental functioning, his skills. Yet, once the depression lifts up, he goes back to see things straight. Hence, depression does behave like a mask for him, but it also does undermine cognitive abilities by preventing his cognitive system from functioning properly.[6]

At the pick of his depressive episodes, Bob lacks the rational skills to adapt to the evidence that he is loved. The point generalizes: being drunk, tired, gaslight, in love, in panic makes us less able to exercise our rationality by slowing our reactions, messing with the firing of our neurons and so on (it's a 'finking', a retreat of one's rationality; Martin 1994). The masks impede the functioning itself, they do not just stop the behavior. That is, some important psychological masking factors impede us to think straight by directly intervening on the source of our ability to reason. And some of these are masks *for beliefs*, beliefs for which we now lack the relevant rational capabilities. Hence, this invalidates (T-2): we can have beliefs for which – *because* of masking factors (e.g. alcohol, depression) – we lack the relevant rational skills. True, Flores and Helton can dig their heels arguing *those* impediments do not count as masks but conditions that rule out beliefs. If so, I am afraid they face the same objection (T-1) faced: denying the status of belief to many attitudes that really look like beliefs.

[6] A non-psychological example. Imagine an opera singer able to sing the most difficult arias: she gets cancer to her vocal cords and she can't produce any singing anymore. Without cancer, she'd be able to sign. In principle she still could sing: in ideal conditions, without cancer, she would sing (Mumford 1998). Hence, it does look like the cancer is a mask, an impediment to the singing, according to what is said above: once the cancer is gone she will sing again. But at the moment the cancer also interferes with the very mechanism that allows her to sing: the cancer erodes her ability because it attacks the mechanism that produces the singing. Hence, the masking factor that does not allow her to perform, does so by damaging her abilities, crippling the source of her skills. This mask is also a capability-eroding condition.

I have a second pressing worry: (T-2) is not really helpful to distinguish belief from other cognitive attitudes such as imagination and acceptance. This concern has to do with the ability of (T-2) to help psychological inquiry in determining when an attitude is a belief. More specifically, the necessary condition for belief listed in T-2 does not offer a clear path for us to separate belief from imagination, acceptance and the rest.

Helton (2018: 3) explains that her "view requires only that subjects be able to revise their beliefs, not that they be inclined to do so." Thus, absent other tools for distinguishing belief vs imagining, we'd be forced to say that a belief could behave perfectly like a piece of imagining, without them being ever distinguishable. Similarly, Flores (2021; MS) refers to subjects' capacities to respond to evidence in some "ideal conditions", suitable for the use of our evidence responsiveness capacities. Ideal conditions may not occur when masking factors are in place. Attitudes that are not updated in the light of evidence even in such ideal conditions do not count as beliefs. Yet if these conditions never materialize themselves, we would not be able to distinguish belief from the rest.

This leads to a related issue: in some ideal condition, even imagination or acceptance may *respond* to the evidence. After playing a game of make believe, the child is able to update its imaginations (e.g. the banana is a phone) according to the evidence. As a result, she stops imagining, and lets go of her pretending (e.g. the banana is not a phone now). It appears that the child had the ability (or capacity) to revise her imagining according to the evidence all along.

To illustrate this further, take the case of religious credence. In some context, I may have the religious credence that the Gods see us. In others, where truth and epistemic values are in place, I may drop that credence all together, only to regain it when I am again in a religious context. Is that credence a belief (that the Gods see me) or a piece of imagining (or another neighboring attitude)? I think this is indetermined based on (T-2). When immersed in a religious context, there are values at play that hinder or undermine any motivation to find the truth. When those drop, rational abilities emerge. However, it is unclear whether those conditions count as *masks that impede beliefs from updating, or conditions that facilitate imagining*. Under (T-2), it is undetermined whether the religious context counted as a mask (allowing the person to irrationally believe the Gods were watching), or as a context that induced the person to *imagine* that the Gods were watching. And yet, being this a topic of

great controversy, this is an issue our definition of belief needs to be able to address in order to inform a mature science of belief. [7]

In sum, invoking masking factors – as Flores does – can be used to explain belief's behavior *only* once we have a clear picture of what a belief is. Only then can we explain what constitutes a mask for belief and what does not. In contrast, it seems that those adopting (T-2) put the horse before the cart, invoking the idea of masking to separate the good vs. the bad cases, namely the cases of imagining vs irrational believing. Unfortunately, this strategy leaves us with little guidance on what to do in the cases I just illustrated.

3. Arguing for Minimalism: evidence in favor and explanatory power

Humans are irrational and their beliefs are at times stubbornly resistant to the evidence. And yet, no matter how epistemically irrational they may be, beliefs' irrationality has a limit. More specifically, humans are not completely indifferent to their own doxastic failings. And the way they react to their beliefs' irrationality is the mark of belief, it is the way we can distinguish belief from other attitudes. Thus, this is the view I will try to defend in the rest of the paper:

> Minimalism: a subject S' cognitive state D is a belief only if S is (at least) minimally doxastically rational with respect to D[8]

[7] Those accepting T-2 could reply that all they are doing is offering one necessary condition for belief: there might be other ways to set belief apart from imagination and the rest. However, we still want our theory of belief to explain what testable predictions we can make about belief vs imagination when it comes to evidence-sensitivity. It seems that (T-2) cannot offer that.

[8] Among the supporters of this view, we likely find Currie (2000), Davidson (1984) and Dennett (1989). Cherniak (1996) talks about 'minimal rationality' but he means something different.

Minimal rationality is a *reactive* form of rationality: S is minimally doxastically[9] rational with respect to D when S will react to D's perceived irrationality by taking measures to reestablish rationality.

As a result, let me restate my view:

> Minimalism: a subject S' cognitive state D is a belief only if – when irrationality is detected – S will make sure that D is in line with other attitudes S has, and do so in a way that would make D doxastically *rational* with respect to other attitudes.

Importantly, Minimalism is not meant to offer a complete picture of belief's behavior given that it focuses only on what happens when belief is irrational and such an irrationality is detected. However, Minimalism is able to show that the type of irrationality- minimization strategies in place for belief are unique, and those strategies differentiate belief from other mental attitudes.

To be clear, the type of irrationality that is at the core of Minimalism is incoherence. Here I focus on *epistemic* incoherence but nothing excludes that this can be broadened to other types of belief's irrationality. Let's see what epistemic incoherence is here. The literature on rationality makes the distinction between substantial vs. structural rationality (Scanlon, 2007; Broome, 2013). Structural rationality (or coherence) has to do with how attitudes relate to one another, whereas substantial rationality requires responding to the reasons one has. Minimalism is interested in beliefs' motivation to conform to norms that roughly look like (formal and non-formal) standards of *structural* rationality. For instance, Modus Ponens says that rationality requires subject S that {S believes that q, if S believes that p & S believes that if p then q}. Detected violations of Modus Ponens will be met with attempts to restore rationality either by coming to believe that q or by ceasing to believe that p (or that if p then q). Similarly, if I believe that p and I believe that my evidence conclusively indicates that non-p, I am violating a coherence constraint. In fact, epistemic structural rationality prohibits the following combinations: {believing that p, while believing you do not have enough reason for

[9] In this paper I only focus on belief's epistemic rationality but doxastic rationality is broader and includes belief's output norms as well. These regulate how belief should interact with action, cause the appropriate behavior and produce the right sort of inferences (Stich 1978: 507).

that belief}; {not believing that p, while believing you do have enough reason for the belief that p}. It also prohibits contradictions: {believing that p, while also believing that p is false}. Importantly, however, the psychological contradictions mentioned by Minimalism go beyond the forms of irrationality we just saw. This is because people may perceive attitudes as being in conflict even if they are in fact structurally rational.[10] Johnson-Laird et al. (2004) explain that people judge consistency based on mental models. If mental models can be created, then the propositions that make them up are consistent. If no model can be created, then they are seen as inconsistent. For instance, it is psychologically impossible to build a model in which "Rome is north of London", "London is north of Paris", "Paris is north of Rome" are all true. Similarly, it is also psychologically difficult to form a model in which one believes one should never vote for Republicans while also knowing her best friend likes the Republican party. As studies indicate, we tend to like those who are liked by the people we like, and dislike those who are disliked by our friends (Gawronski, Walther, & Blank, 2005). Therefore, the idea that someone we love is into people we do not like may be perceived as a contradiction, even if this is no logical inconsistency (Gawronski & Brannon 2019: 94-95).

Minimal rationality requires that, when irrational combinations come into view, we are motivated to solve the tension among those attitudes. This motivation shows up in three ways: (1) a sense of distress when irrationality is detected; (2) active psychological measures are taken for reestablish coherence among the contrasting attitudes; (3) these measures do not necessarily bring beliefs to be epistemically rational (i.e. to align with the epistemic reasons one has).

Note that this rational baseline can be guaranteed through an automatic process: complex reflective cognitions are not necessary for detecting irrationality and satisfying coherence requirements. Of course, that does not guarantee that our beliefs or minds are in fact overall coherent. The motivation here is to eliminate perceived incoherence, not actual incoherence.

[10] "Although Festinger's (1957) original definition puts a strong emphasis on logical relations between cognitive elements […] he explicitly acknowledged the role of cultural mores, opinions, and personal experiences as important determinants of perceived (in)consistency. In this sense, cognitive (in)consistency can be said to describe psycho-logical (rather than strictly logical) relations between cognitive elements" Gawronski & Brannon (2019: 94).

As we will now see, evidence for Minimalism comes from various, alternative sources. In particular, Minimalism squares nicely with various psychological theories and findings. It also matches with some of our key intuitions about beliefs such as the Moore's paradox. Finally, it provides a clear path to distinguish beliefs from other neighboring attitudes such as imagination and belief, while also offering a way to understand the nature of delusions. That is, Minimalism lays out a necessary condition for an attitude to be belief. Though not sufficient to completely define belief, this condition is fundamental for cognitive science's ability to make testable predictions about how epistemically irrational beliefs behave.

3.1 Support for Minimalism: psychology and cognitive science

Though Minimalism is presented as a theory concerning belief, it is meant to fit within a broader view of the mind. Support for Minimalism can be found in cognitive consistency theories in social psychology (Heider 1946, 1958; Newcomb 1953, Osgood and Tannenbaum 1955) and in recent observations about how beliefs tend to behave (Porot and Mandelbaum 2020).

Cognitive consistency theories in social psychology claim that humans are driven by the desire to have coherent attitudes. Recently models of coherence-based reasoning have been used to explain decisions in widely different domains, from medical, moral and legal reasoning (Holyoak & Simon 1999; Simon 2004; Holyoak & Powell 2016) to evidence evaluation (Carlson & Russo 2001; Glöckner, Tilmann & Schindler 2010). These models describe decision making processes in which complex decisions result from a bidirectional unconscious mechanism that reinforce coherence among attitudes (Read & Simon 2012; Simon & Holyoak 2002; Russo et al. 2008). On this view, coherence-seeking is thus the driving force behind attitudes formation.

Among all cognitive consistency theories, it is dissonance theory that has become dominant. Such a theory points out that spotting harboring conflicting attitudes of a certain kind hurts and, when that happens, we are pushed to change our attitudes to eliminate the conflict (McGregor, Newby-Clark, & Zanna 1999). This phenomenon is called "cognitive dissonance"

(Festinger, 1957; Harmon-Jones et al 2009; Van Veen et al. 2007). As Festinger explained, "[c]ognitive dissonance can be seen as an antecedent condition which leads to activity oriented toward dissonance reduction just as hunger leads toward activity oriented toward hunger reduction". Though it is recognized that we feel different degree of dissonance depending on how much we care about the incompatible cognitions (Nohlen et al. 2016), many also believe that cognitive dissonance is a universal phenomenon, driving inconsistency-reduction in all situations (Proulx, Inzlicht, & Harmon-Jones, 2012).

Initially, cognitive dissonance was seen as a phenomenon that generally emerged whenever there was a detected incompatibility between *values and actions*. Thus, agents are likely to decrease their view of others when they are forced to show hostile behavior against them (Davis and Jones 1960; Glass 1964). This phenomenon is also present in children and non-human animals (Aronson and Carlsmith, 1963; Friedrich & Zentall, 2004; Egan, et al. 2007).[11] For instance, results indicate that children devalue a toy if they are told they can't play with it (Aronson and Carlsmith, 1963). Lately, more focus has been devoted to the incompatibility between beliefs (Burris, Harmon-Jones, & Tarpley, 1997; Gawronski, Ye, Rydell, & De Houwer, 2014; Gawronski & Brannon 2019). Minimalism as presented here is a theory about beliefs that falls in line with the findings of psychology. These findings indicate that detected irrationality involving beliefs brings about several dissonance-reduction strategies. We will turn to these now as they can be used to track minimal rationality in irrational beliefs.

3.1.1 Signs of Minimal Rationality in Irrational Beliefs

Sitting within the framework of coherence-based theories, and dissonance theory in particular, Minimalism takes the following important aspects of our cognition to be the telling signs of minimal rationality for belief. That means that attitudes that follow (some or most of) these patterns would count as minimally rational beliefs. These attitudes are beliefs because they retain a level of rationality even when they behave mostly irrationally.

[11] This suggests that dissonance-detection and reduction may not always require metacognitive "high-level capacities" (Egan et al. 2007: 982).

- Selective exposure: there is substantial evidence that people avoid being exposed to information that may clash with pre-existing beliefs (Smith, Fabrigar, & Norris, 2008), and look for information that squares with what they already take to be true (Lord et al 1979; Brannon, Tagler & Eagly, 2007). For instance, people with low self- esteem look for and have a preference for information consistent with their unflattering view of themselves, even though that information comes in the form of harsh, negative feedback (Swann, Wenzlaff, Krull, & Pelham, 1992).
- Belief's perseverance: some studies show that in some cases affirming counter evidence to one's beliefs actually increases the strengths in those beliefs. This is considered a form of anti-Bayesian updating (Plous 1991; Liberman & Chaiken 1992), and operates when the beliefs involved are particularly important to us (e.g. beliefs about the self) or are consonant with a number of other beliefs.
- Rationalization & confabulation: within the cognitive dissonance framework it is well established that, to avoid the discomfort produced by dissonance when dissonance can't be easily eliminated, a subject may try to add new cognitive, attitude-consistent elements. Rationalizations and confabulations are used to solve the conflict between attitudes and new (confabulated) attitudes are created to avoid conflicting cognitions. This type of rationalization – the creation of false and unsupported reasons for one's choices and attitudes – often occurs when subjects are asked to explain and justify behavior or attitudes they hold and that are the result of unconscious processes (Nisbett & Wilson, 1977). Hence, confabulation likely helps us avoid the psychological burden of conflicting cognitions (Carruthers 2011: 356-365; Bergamaschi Ganapini 2020a).
- Fragmentation of the mind: another strategy to deal with conflicting attitudes is to fragment the mind so as to isolate conflicting cognition in separate compartments (Elga & Rayo, 2022). This fragmentation explains why people have contradictory beliefs while also striving for coherence. Beliefs in the same compartment are activated at the same time and so they tend to be coherent. Bendana and Mandelbaum (2021: XX) explain that fragmentation "allows for a belief system riddled with inconsistencies, as long as those inconsistencies are sequestered from one another in

> separate fragments. […] across fragments, consistency (and simplicity) is not maintained, and representational redundancy and inconsistency are expected."
>
> Attitudes in the same fragment come into view at the same time and once within the same fragment, inconsistencies are "automatically resolved": "if two fragments that contain inconsistent information are coactive they will be rendered consistent". As a result, Bendana and Mandelbaum explain that "[e]ach fragment is internally consistent and contains no redundant representations." That is, (Bendana and Mandelbaum, 2021). Such a push is not felt when tokens of incoherent beliefs are not activated at the same time as they may belong to different compartments.

These are the telling signs of minimal rationality. That is, showing some or most of these types of behavior, especially after incoherence and irrationality are detected, is a sign that an attitude is at least doxastically minimally rational and thus is a belief. In contrast, no motivation or attempt to solve incoherence and irrationality when detected is a sign that the attitude is not minimally rational and therefore not a belief.

3.2 Believing, yes …but not at will

Minimalism's explanatory power extends to other commonly observed aspects of belief. In particular, one key feature of belief, that also arguably distinguishes belief from other attitudes such as supposition or imagination, is that we cannot believe at will (Williams, 1973). This phrase has acquired different meanings in the literature, but here I focus on a specific aspect of belief, namely that one cannot *consciously* form a belief as a result of one's desire to do so. That is, desiring to have a belief may *unconsciously* bring me to form that belief. However, I cannot use that desire as a reason to have that belief in deliberation (Shah, 2003; Steglich-Petersen 2006). Imagine a college student: she wants to get good grades to feel good about herself but she struggles to find the time and motivation to study. She may realize that simply believing she got good grades in her finals would make her happy, however she cannot use that as a way to convince herself to believe, unless she also has sufficient evidence to show that she did in fact get good grades. Her desire alone cannot be

used as a reason for belief in doxastic deliberation (Bergamaschi Ganapini 2020b).[12] That is a psychological impossibility widely discussed in the literature (McHugh 2013; Owens 2003; Shah 2003; Sullivan- Bissett 2018) and Minimalism can nicely explain it. Minimalism tells us that we are disposed to eliminate perceived inconsistency, moved by the fact that those produce a feeling of displeasure (i.e. cognitive dissonance). Parallely, this psychological fact can also explain why we cannot bring ourselves to *create* these inconsistencies as a result of a voluntary act of deliberation. To illustrate, let's come back to our college student. She could not find the time nor the motivation to bring herself to study for the finals scheduled for the next day. She has little evidence that she will do well. She'd be happier thinking that she will get good grades, but if that practical consideration were to motivate her in deliberation, she would be finding herself believing she will get good grades while also believing - in full awareness - that her belief is based only on some pragmatic consideration and thus it is epistemically irrational.[13] That is what Minimalism predicts cannot happen: perception of irrationality brings avoidance and pain, thus we cannot voluntarily make our beliefs to be overtly irrational in that way. In conclusion then, Minimalism fully explains why we are not able to believe at will.

3.3 The Moore's Paradox

Though the main focus of this paper has been squaring Minimalism with recent findings in psychology and cognitive science, Minimalism also matches with some well-known intuitive data as well. In particular, the Moore's Paradox, a widely discussed paradox about belief, can be accounted for by using Minimalism. The paradox emerges from noticing the absurdity of uttering sentences like the following:

(P) I went to the pictures last Tuesday, but I don't believe that I did (Moore 1962)

[12] In the metaethical literature this is called a 'wrong kind of a reason' to believe. Wrong kind of reason cannot be the *reasons for which* we believe something.

[13] For an explanation of why epistemic and not pragmatic rationality is dominant in doxastic deliberation see, for instance, Bergamaschi Ganapini (2020b) and Sullivan- Bissett (2018).

Though P can be true, it sounds absurd. Notice that it is completely fine to describe someone else as being gone to the movie and not believing it. The absurdity emerges only when P is a first-person utterance self-describing oneself as being gone to the movie *and* not believing it. There are several other versions of P (Sorensen 1988) and there have been many attempts to explain why P sounds so weird. The debate on the Moore Paradox is incredibly vast and I can only hint at it here. One thing to notice is that the Paradox seems to be due to some kind of contradiction harbored in statements like P (DeRose, 2009). Though the conjuncts are not building a formal contradiction, some scholars have pointed out that the absurdity lies in the fact that saying "I went to the pictures last Tuesday" is tantamount to expressing your belief that you went to the pictures last Tuesday. If this is right, then the absurdity emerges out of the fact that one is overtly showing a tension within one's attitudes: the belief that I lack an attitude that I am also, at the same time, expressing. Though our mind is probably riddled with these kinds of contradictions, the fact that one is, in full awareness, pointing to their contradictory beliefs sounds absurd. And Minimalism can make sense of that absurdity: Minimalism predicts that once such a contradiction emerges one takes immediate steps to solve it. It is thus very surprising to hear expressing contradictory beliefs *in full view* because that is *not* how believers actually deal with irrationalities (according to Minimalism). Although it is compatible with Minimalism that one might not be able to get rid of contradictions, seeing someone straight-out pointing at such a self-contradiction without any caveat or explanation, is in tension with our experience concerning belief and believers. As a result, it sounds psychologically very implausible to us. Hence, according to Minimalism, it is not surprising that we find Moorean sentences of that kind quite absurd.

4. Minimalism: explaining imagining and other borderlines cases

Minimalism is well suited to explain how belief is different from other neighboring attitudes and to offer a way to test if an attitude is a belief or not. Let's focus on the distinction between belief and imagining (still what I say here holds for other attitudes such as accepting or

supposing as well)[14]. Traditionally it is said that imagination differs from belief because the latter but not the former is evidence-responsive. In this paper, I showed that evidence-responsiveness is not a useful tool to differentiate belief from the rest in some of the hard cases. Now I propose to use Minimalism as a guide instead. According to Minimalism, perceived disunity affects belief differently than imagination: imaginations and the like are governed by a decoupling mechanism that produces fragmentation and disunity (Leslie, 1987). For imagination disunity is a feature, not a bug. That is, imagination is structured in (usually) sealed compartments (or worlds), and no coherence is required across those compartments. As a result, we can imagine different, competing scenarios to be true at the same time. In contrast, we cannot, in full awareness, *believe* different, competing scenarios to be true at the same time. Similarly, we cannot anchor belief to context: I cannot say that I believe the Earth is flat *at mass* while also sincerely rejecting that very belief because now I am *at work*. In contrast, we can easily conceive of a director who is filming two different movies at the same time while imaging different things depending on which movie-set she is at. Her attitudes will switch simply based on the context she is in, and she can – in full awareness – track what imaginings belong to each context. For belief, perceived fragmentation and compartmentalization give rise to cognitive dissonance and a motivation to get rid of incoherence *across contexts*. This particular type of motivation, namely the motivation to establish an apparent *doxastic* coherence, is the mark of belief. Hence, if we see a subject happily harboring conflicting cognitive attitudes about *p* we have reasons for taking her not to believe that p but to imagine it. In contrast, we are at a presence of a belief if we see our subject trying to amend her epistemically irrational attitude (Bergamaschi Ganapini 2020).

4.1. Delusions: are they beliefs?

Delusions are described as cognitive attitudes that are mostly resistant to change even in light of overwhelming counterevidence (APA, 2013). They are commonly associated with

[14] Dennett's "opinion" (1978) and Sperber's "reflective belief" (1996) can be seen as forms of acceptance.

severe mental illness. Because of how they seem to be evidence-impenetrable, delusions behave differently than ordinary factual beliefs. As we saw above these beliefs are characterized as being mostly evidence-sensitive: they change in relation to the evidence the subject has. This consideration has brought many to doubt that delusions can be beliefs (e.g. Currie & Ravenscroft 2002; Schwitzgebel 2012).

Contrary to this anti-doxasticist position, some researchers have argued in favor of the claim that delusions are beliefs (Bortolotti 2012). They point out that even ordinary beliefs are evidence-resistant sometimes. Eve ordinary beliefs can be irrational: absurd political beliefs, overly-optimistic beliefs are all examples of attitudes we hold in the face of overwhelming contrary evidence (Bortolotti 2010, 2018; Bell et al 2021). Hence, if we are prepared to call those beliefs, why not extend the same courtesy to delusions as well?

The debate between doxasticist and anti-doxasticist has been ranging for a few years now and I believe Minimalism can offer a reasonable way out of the gridlock. Minimalism proposes this test for belief: to know whether an attitude is a belief, see how the attitude and the subject react to detected irrationality and internal incoherence (Bergstein 2008; Currie 2000). If you see discomfort and coping strategies to avoid conflict, that's a belief. As such, Minimalism has a way to settle the debates between doxasticism and antidoxasticism about delusions. For starters, to know whether delusions are beliefs one would need to look at specific cases. As we will see momentarily, there are many different types of delusions and so it is difficult to generalize and say whether or not *all* delusions are beliefs.

In the literature on delusions, we see many cases of clinical delusions in which subjects, when faced with obvious counter-evidence for their delusions, would either give up on their delusion (after a while) or feel distress and engage in confabulations and evidence-avoidance (Amanda et al 2010). This latter behavior is found consistently both in clinical monothematic delusions (i.e. delusions that are self-contained and revolve around one specific topic) and in polythematic delusions (i.e. delusions that cover diverse topics and tend to expand).

Frazer and Roberts (1994: 557) report of a woman suffering from Capgras delusion about her son: she insisted that her son had been replaced by an impostor. When it was pointed out to her that her son and the alleged impostor looked very similar, she confabulated that the impostor "had different-coloured eyes, was not as big and brawny as her son, and that

her real son would not kiss her". Being confronted with her inability to give up her delusion, the woman does not simply accept the contradiction but tries to confabulate it away.

Here's another example:

"Being dead butted up against the so-called evidence of being alive, and so I grew to avoid that evidence because proof was not a comfort; instead, it pointed to my insanity". (Wang 2019: 157) This subject shows that they have an impaired capacity to rationally update their belief based on the evidence, but they are still feeling the cognitive dissonance of their clashing attitudes. Similarly, Ramchandra et al (1996: 349) explain that "[w]hat is much more common in patients with anosognosia is a tendency to come up with all sorts of rationalizations to explain why the arm doesn't move; they don't usually say that they can actually see their arm moving." Hence, their delusion concerning the status of their limbs is sustained by using a confabulatory process which we saw is one of the typical responses to cognitive dissonance and a sign of minimal doxastic rationality.

Whereas delusions such as Capgras delusion focus on one specific topic, patients with schizophrenia tend to have polythematic delusions. These are fairly elaborate sets of delusions which usually emerge out of anomalous sensory or somatic experiences. Describing the experience of one of these patients, Upthegrove and Allan (2018: 6) write: "Andrew's belief arises secondary to his array of tactile, somatic and auditory hallucinations. He uses extensive and increasing information to reinforce his firm belief and discounts evidence that challenges his conviction." A complex array of other delusions is now formed in the mind of the patient, which may provide them with a sense of coherence. In this case, the delusions are used as explanations to make sense and give meaning to odd and disturbing experiences (Mishara & Corlett, 2009: 531). Again, here we find one of key traits of *minimally* doxastically rational attitudes: they are formed to give meaning and coherence to other attitudes, and when they are challenged by counter-evidence, they may give rise to confabulatory explanations. If that's the case, then we have some initial reason to think these delusions are in fact beliefs.

In conclusion, Minimalism sides with doxasticism in claiming that epistemic irrationality may not be enough to disqualified delusions from being beliefs.[15] Yet, no broad generalization about delusions can be made without looking at the specific cases and checking for signs of minimal doxastic rationality.

4.2. Conspiracy theories and fake news

Do people really believe the conspiracy theories and fake news they sincerely share? This question mirrors the debate that we have just covered between doxasticism and anti-doxasticism about delusion: some of these conspiracy theories are so blatantly absurd that theorists have started to question whether people believe them for real or whether they are in the grip of some sort of fantasy.[16] As in the case of delusions, it is difficult to generalize and make broad statements without looking at specific cases. One thing worth mentioning is that in many cases those sharing fake stories do not seem to care that those stories are not justified and they are completely indifferent vis-à-vis their own epistemic status as believers. Similarly, conspiracy theories shown to be completely unjustified are endorsed nonetheless, and when hard-pressed to justify their claims, many conspiracy theorists are happy to admit they have little evidence for them. These absurd stories also often openly admit contradictions: conspiracy theorists go around saying both that Lady Diana was murdered and that she is still alive, hiding somewhere (Muirhead and Rosenblum 2019). That is, there is some evidence that those who subscribe to conspiracy theories are disposed to assent to their contradictions too (Wood et al., 2012; Lewandowsky et al. 2016). When it comes to political conspiracy theorizing, we see a lot of people approaching it in a ironic way, as a

[15] There may be other reasons for thinking that some delusions are not beliefs. For instance, some delusions are behaviorally encapsulated and reliably fail to produce any suitable inference. In this case and especially if they also lack any sign of minimal rationality, their inferential and practical inertness may be a reason for thinking those particular delusions are not beliefs. Also note that Currie (2002:160) assumes an anti-doxastic position about delusions because he thinks deluded subjects are mostly indifferent to contradictions (and thus show not minimal rationality; Currie & Jureidini 2001). See Bayne & Pacherie (2005) for a reply to this.

[16] There are studies investigating whether sharing fake news is a form of expressive behavior (for a discussion of this see Levy 2022) while others have argued that sharing fake stories (especially when blatantly absurd) is really a matter of signaling one's social positioning and group membership (Bergamaschi Ganapini 2021).

mixture of fun and fiction combined with intriguing speculations which often will enrage their political opponents (Birchall 2006). Blatant contradictions, manifestly insufficient evidence, little interest in searching for the truth are signs that minimal rationality is at times overtly violated when fake news is concerned, especially when those stories are openly absurd. In some cases, those sharing these stories do not seem motivated to resolve those contradictions, they show little sign of cognitive dissonance or desire to address (or even to cover) their irrationalities. When that occurs, it is likely that even those who sincerely share fake stories do not actually believe them to be true. For all the other cases, a more careful analysis and observation may be needed.

References:

Barnier A. et al. (2010). A Stranger in the Looking Glass: Developing and Challenging a Hypnotic Mirrored-Self Misidentification Delusion, *International Journal of Clinical and Experimental Hypnosis*, 59:1, 1-26

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

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.

Bayne, T., & Pacherie, E. (2005). In defence of the doxastic account of delusions. Mind and Language, 2, 163–188.

Bell, V., Raihani, N., & Wilkinson, S. (2021). Derationalizing Delusions. *Clinical psychological science : a journal of the Association for Psychological Science*, *9*(1), 24–37

Bendana, J & Mandelbaum, E (2021). The Fragmentation of Belief. In Cristina Borgoni, Dirk Kindermann & Andrea Onofri (eds.), The Fragmented Mind. Oxford, UK:

Bergamaschi Ganapini, M. (2020) Belief's Minimal Rationality Philosophical Studies 177 (11): 3263-3282
Bergamaschi Ganapini, M. (2020a) Confabulating Reasons, Topoi 39, 189–201
Bergamaschi Ganapini, M. (2020b). No epistemic norm or aim needed. Episteme, 1– 16
Bergamaschi Ganapini, M. (2021) The Signaling Function of Sharing Fake Stories. *Mind & Language*

Bergstein, M., Weizman, A., & Solomon, Z. (2008). Sense of coherence among delusional patients: Prediction of remission and risk of relapse. *Comprehensive Psychiatry, 49*(3), 288–296.

Birchall, C. 2006. *Knowledge Goes Pop: From Conspiracy Theory to Gossip*. Oxford: Berg.

Bird A. Dispositions and antidotes, Philosophical Quarterly, 1998, vol. 48 (pg. 227-34)

Bortolotti, L. (2010). *Delusions and Other Irrational Beliefs*, Oxford: Oxford University Press.
— (2012) Doxasticism About Delusions. *Neuroethics*, *5*(1), 39–53.
—— (2018). Delusions and three myths of irrational belief. In L. Bortolotti (Ed.), *Delusions in context.* Palgrave Macmillan.

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

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.

Burris, C. T., Harmon-Jones, E., & Tarpley, W. R. (1997). "By Faith Alone": Religious agitation and cognitive dissonance. *Basic and Applied Social Psychology, 19*(1), 17–31.

Carlson, K. A., & Russo, J. E. (2001). Biased interpretation of evidence by mock jurors. Journal of Experimental Psychology: Applied, 7(2), 91–103

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

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

Currie, G. 2000: Imagination, delusion and hallucinations. In M. Coltheart and
M. Davies (eds.), Pathologies of Belief. Blackwell, 167–182.
Currie, G. and Jureidini, J. 2001: Delusion, rationality, empathy. Philosophy, Psychiatry & Psychology, 8/2,3, 159–62.
Currie, G and I Ravenscroft. *Recreative Minds*. Oxford: Clarendon Press, 2002.

Cushman F. (2020) Rationalization is rational. Behavioral and Brain Sciences 43, e28: 1–59. doi:10.1017/ S0140525X19001730

Currie, G. and Jureidini, J. (2002): Delusion, rationality, empathy. Philosophy, Psychiatry & Psychology, 8/2, 159–62.
Davidson, D. (1984). *Inquiries Into Truth And Interpretation*. Oxford University Press.

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

DeRose, K. (2009). *The case for contextualism*. Oxford: Clarendon.

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.

Elga, A. Fragmentation and logical omniscience. *Noûs*. 2022; 56: 716– 741.

Fara, M. (2008). Masked Abilities and Compatibilism. *Mind*, *117*(468), 843–865.

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

Flores, C. (2021), "Delusional Evidence-Responsiveness," Synthese
Flores C (MS), Resistant Beliefs, Responsive Believers

Frazer SJ, Roberts JM (1994) Three cases of Capgras' syndrome. Br J Psychiatry 164:557–559

Gawronski, B., & Brannon, S. M. (2019). What is cognitive consistency, and why does it matter? In E. Harmon-Jones (Ed.), *Cognitive dissonance: Reexamining a pivotal theory in psychology* (pp. 91–116). American Psychological Association.

Gawronski, B., Ye, Y., Rydell, R. J., & De Houwer, J. (2014). Formation, representation, and activation of contextualized attitudes. *Journal of Experimental Social Psychology, 54,* 188–203.

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: Cognitive balance affects the encoding of social information. *Journal ofExperimental Sodal Psycholo9Y, 41,* 618-626.

Glöckner, A & B, Tilmann & Schindler, N. (2010). Coherence Shifts in Probabilistic Inference Tasks. Journal of Behavioral Decision Making. 23. 439 - 462

Glöckner, A., Hilbig, B. E. & Jekel, M. (2014) What is adaptive about adaptive decision making? A parallel constraint satisfaction account. Cognition 133(3):641–66

Glüer, K., & Wikforss, Å. (2013). Against belief normativity. In T. Chan (Ed.), *The aim of belief* (pp. 80– 99). Oxford University Press.

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.

Heider, F. 1946 "Attitudes and Cognitive Organizations." *Journal of Psychology* 21:107–112.
——1958 *The Psychology of Interpersonal Relations*. Hillsdale, N.J.: Erlbaum.

Helton, G. (2019), If You Can't Change What You Believe, You Don't Believe It. *Noûs*.
Holyoak, K. J. & Powell, D. (2016) Deontological coherence: A framework for commonsense moral reasoning. Psychological Bulletin 142:1179–1203.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128*(1), 3–31.

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review, Ill,* 640-661.

Levy N., *Bad Beliefs: Why They Happen to Good People*, Oxford University Press, 2022,

Lewandowsky, S., Cook, J., & Lloyd, E. (2016). The “Alice in Wonderland” mechanics of the rejection of (climate) science: Simulating coherence by conspiracism. *Synthese*, *195*, 175–196.

Liberman, A., & Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin, 18*(6), 669–679.

Lord, C. G., Ross, L. & Lepper, M. R. (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

Mandelbaum, Eric & Porot, Nicolas (forthcoming). The Science of Belief: A Progress Report (Expanded Reprint). In Joseph Summer Julien Musolino (ed.), *The Science of Beliefs: A Multidisciplinary Approach. Cambridge University Press.* Cambridge, UK:

McGregor, I., Newby-Clark, I.R., & Zanna, M.P. (1999). "Remembering" dissonance: Simultaneous accessibility of inconsistent cognitive elements moderates epistemic discomfort. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (pp. 325-353). Washington, DC: American Psycho- logical Association.

McHugh, Conor (2012) 'The Truth Norm of Belief', Pacific Philosophical Quarterly 93(1), 8-30.

McHugh, C. (2013). The illusion of exclusivity. European Journal of Philosophy, 23, 1117– 1136.

McHugh, Conor (2014) 'Fitting Belief', Proceedings of the Aristotelian Society 114(2), 167-187.

Mishara, A. L., & Corlett, P. (2009). Are delusions biologically adaptive? Salvaging the doxastic shear pin. *Behavioral and Brain Sciences, 32*, 530–531

Moore, G.E. (1962). *Commonplace Book: 1919–1953*. London: George Allen & Unwin.

Muirhead, R. and Rosenblum, N.L. (2019) *A Lot of People are Saying: The New Conspiracism and the Assault on Democracy*. Princeton, NJ: Princeton University Press.

Mumford S. , Dispositions, 1998 OxfordOxford University Press

Newcomb, T.M. (1953) An Approach to the Study of Communicative Acts. Psychological Review, 60, 393-404.

Nisbett RE & Wilson TD (1977). Telling more than we can know: verbal reports on mental processes. *Psychol Rev* 84:231–259

Nohlen, H. U., van HarrevelcL F., Rotteveel, M., Barends, A. J., & Larsen, J. T. (2016). Affective responses to ambivalence are context-dependent: A facial EMG study on the role of inconsistency and evaluative context in shaping affective responses to ambivalence. *Journal of Experimental Social Psychology, 65,* 42-51.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review, 62*(1), 42–55.

Owens, D. (2003). Does belief have an aim? Philosophical Studies, 115, 283– 305.

Plous, S. (1991), Biases in the Assimilation of Technological Breakdowns: Do Accidents Make Us Safer?. Journal of Applied Social Psychology, 21: 1058-1082.

Proulx, T., Inzlicht, M., & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in cognitive sciences*, *16*(5), 285–291

Ramachandran, Vilayanur S.; Levi, L.; Stone, L.; Rogers-Ramachandran, D.; and others.

Illusions of body image: What they reveal about human nature. In *The mind-brain continuum: Sensory processes*. Edited by Rodolfo Riascos Llinas, Patricia Smith Churchland, et al. Cambridge, MA: MIT Press, 1996. 29-60

Read, S. J. & Simon, D. (2012) Parallel constraint satisfaction as a mechanism for cognitive consistency. In: Cognitive consistency: A fundamental principle in social cognition, ed. Gawronski, B. & Strack, F., pp. 66–86. Guilford Press.

Russo, J. E., Carlson, K. A., Meloy, M. G. & Yong, K. (2008) The goal of consistency as a cause of information distortion. Journal of Experimental Psychology: General 137(3):456–70.

Shah, N. (2003). How truth governs belief. Philosophical Review, 112, 447– 482.

Simon, D. (2004) A third view of the black box: Cognitive coherence in legal decision making. University of Chicago Law Review 71:511–86.

Simon, D. & Holyoak, K. J. (2002) Structural dynamics of cognition: From consistency theories to constraint satisfaction. Personality and Social Psychology Review 6:283–94.

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

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

Schwitzgebel, E. 'Mad Belief?' *Neuroethics* **5**.1 (2012): 13–7.

Sorensen, R. A. 1992. "Moore's Paradox". In J. Dancy and E. Sosa, eds., A Companion to Epistemology. Oxford: Basil 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.

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

Steglish- Petersen, A. (2006). The aim of belief: No norm needed. Philosophical Quarterly, 56, 500– 516

Stein E, *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*, Oxford, Clarendon Press, 1996

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

Sullivan- Bissett, E. (2018). Explaining doxastic transparency: Aim, norm, or function? Synthese, 195(8), 3453– 3476.

Swann, W. B., Wenzlaff, R. M., Krull, D. S., & Pelham, B. W. (1992). Allure of negative feedback: Self-verification strivings among depressed persons. *Journal of Abnormal Psychology, 101*(2), 293–306

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

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

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

Upthegrove, R., & Allan, S. (2018). Delusional beliefs in the clinical context. In *Delusions in Context* (pp. 1-34). Springer. https://doi.org/10.1007/978-3-319-97202-2_1

Viedge, N. (2018) Defending Evidence-Resistant Beliefs. *Pacific Philosophical Quarterly*, 99: 517– 537.

Wang, E. W. (2019). The collected Schizophrenias. Minneapolis: Graywolf Press.

Williams B. (1973). Deciding to Believe, in *Problems of the Self.* Cambridge: Cambridge University Press

Wood, M. J., Douglas, K. M., & Sutton, R. M. (2012). Dead and Alive: Beliefs in Contradictory Conspiracy Theories. *Social Psychological and Personality Science*, *3*(6), 767–773