Rationalism

This is a pen-ultimate draft please cite the published version, in the Routledge Handbook of the Philosophy of Delusion *(ed. Ema Sullivan-Bissett)*

Jakob Ohlhorst – VU Amsterdam – j.ohlhorst2@vu.nl

Abstract:

This chapter introduces the rationalist model of delusions. It begins by presenting John Campbell's seminal proposal that delusions are caused top-down by pathological Wittgensteinian framework or hinge beliefs. After presenting Campbell's rationalist account of delusions, the chapter raises and examines prominent objections by Tim Bayne & Elisabeth Pacherie as well as by Tim Thornton. The former make an important distinction between the aetiological top-down cognitive part and the epistemological rationalist framework part of Campbell's account. The thesis that delusions are caused top-down by pathological beliefs is not equivalent to the thesis that delusions are Wittgensteinian framework certainties. This chapter endorses this distinction. While the arguments against the top-down aspect are found to be more convincing – which motivates popular two-factor theories – the rationalist framework model of delusions is defended as an epistemologically promising account of delusion. Finally, this chapter examines a range of developments that have been made taking Campbell's rationalism as a starting point. Most prominently the section examines Eilan's early development of rationalism and Rhodes and Gipps' proposal that delusions are not framework propositions but a product of their absence.

Introduction

Rationalism about delusions is the theory that the root of delusions lies in a *cognitive* disturbance. The idea is quite simple: delusions are disturbed thoughts and consequently the delusional patient's *thinking* must be the source of the disturbance. Rationalism contrasts with empiricism about delusions (see Chapter 27) which posits that delusions are the products of a disturbed *experience*. In a way, rationalism and empiricism about delusions are the extension of traditional rationalism and empiricism – the extension from theories about the sources of our knowledge into theories about the sources of our delusions.

Rationalism was originally proposed by John Campbell in his 'Rationality, Meaning, and the Analysis of Delusion' (2001). The distinction between rationalist and empiricist accounts of delusions has proven highly influential in the debate. Following Campbell's terminology, on the empiricist account, delusions are also described as 'bottom-up' where the experience's content is the cause of the delusion. On rationalist accounts, delusions are 'top-down' and the delusional cognitive content informs the content of the patient's disturbing experience. (Campbell, 2001, pp. 95–96; Hohwy, 2004)

In this chapter, I will first give an account of Campbell's argument for rationalism about delusions as well as his theory of delusion. Second, I will present the most prominent challenges to rationalism about delusions and respond on behalf of the rationalist. Third, I will present developments of the rationalist account and how they may address these challenges. Rationalism is a niche position in the epistemology of

delusions, but due to its alignment with Wittgensteinian hinge epistemology, it has profited from the latter's growing popularity in recent years.

Campbell's Argument

Campbell (2001, p. 89) presents delusions in light of the problem of radical interpretation. (Quine, 1960; Davidson, 1968) Namely, if we want to meaningfully ascribe any mental state – beliefs, hopes, emotions – to some agent, then we are forced to assume that they are rational. We cannot meaningfully describe a person as believing something if we do not think that they are subject to rationality, i.e. that they are not constrained by coherence and some kind of reasons.[1]

Pathological delusions *are ascribed as beliefs*, but at the same time they are also described as un-understandable, beyond rationality and constituting a *breakdown* of rationality. This gives us two options: Either we treat a person who exhibits delusions as absolutely arational, 'crazy', and lacking what we would be willing to describe a mental life with beliefs, hopes, and desires; or we grant that the delusional person has meaningful beliefs, hopes and desires, i.e. a mental life, we grant that they exhibit some kind of rationality, and we take delusions to be meaningful doxastic states.

Like many others, Campbell (2001, p. 89) opts for the latter option – also delusional patients have a describable mental life. But this then raises the question: How can a person exhibiting such strange beliefs and being insensitive to all sorts of reasons possess any rationality? The delusional patient is not understandable by the ordinary routes that we use to understand each other, they cannot give convincing reasons

[1] Bortolotti (2005) attacks this Davidsonian view of mental life, arguing that delusions are beliefs, even though the patient is patently irrational.

for their delusions and they fail to appreciate the force of reasons we would give against their delusions – hence the appearance of a- or irrationality.

Campbell suggests that there are two ways how the delusional patient can acquire these apparently arational and un-understandable beliefs: They come 'bottom-up' and are the product of strange, un-understandable, and disturbing experiences on whose basis the delusional beliefs are formed because the delusional content is the only available explanation. (Stone and Young, 1997) Campbell (2001, p. 89) calls this approach *empiricism*. He criticises that strange experiences are not sufficient to rationally explain why the patient would for instance hold such strange delusions like *I am dead* with Cotard's syndrome. After all, rationally speaking, I can never have evidence for my own death because, once I am dead, I am unable to have beliefs and possess evidence. The most reasonable response to an unsettling and strange experience would arguably be to recognise that it is pathological.

As an alternative solution, he proposes, that delusions are produced 'top-down' by our belief system. That is, what is at the heart of a delusion is the formation of strange beliefs which are the delusions, and these subsequently influence and shape the patient's experience. Campbell (2001, p. 89) calls this view *rationalism* about delusions.

Many authors are convinced by Campbell's argument against empiricism but less by his positive rationalist proposal. Consequently they opt for a third option integrating rationalism and empiricism. (Hohwy, 2004, p. 66) Such two-factor theories argue that strange and unsettling experiences are interpreted by some defective top-down cognitive mechanism to form a delusion. I will examine one such two-factor theory that shares a considerable overlap with Campbell's rationalism further below. But

most two-factor theories do not involve a rationalist element – instead their top-down mechanism is a bias or a cognitive error, i.e. a kind of irrationality. (See Chapter 30)

How can we then make sense of delusions rationalistically, i.e. as a top-down phenomenon? Coming back to the problem of radical interpretation: If someone uses language in an apparently irrational and un-understandable way, while we nevertheless assume that this person is rational and has a mental life that can in principle be interpreted, then the only reasonable inference is that this person uses language differently from us. That is, the delusional patient is not arational and insensitive to reasons, but rather they mean something altogether different than what we understand from their utterances. For instance, 'I am dead' in the case of Cotard syndrome is not incompatible with the person still holding beliefs or breathing – the words 'I', 'am', and 'dead' work differently in this case. It is worth noting that Campbell only focuses on monothematic delusions, i.e. the Cotard and Capgras syndromes.

But how would a delusional patient come to such a profound transformation in their language? Campbell (2001, p. 96) proposes that Ludwig Wittgenstein's *On Certainty* (1969) can shed light on this shift. Namely, Campbell suggests that delusions are the product of a shift in our 'framework', or *hinge*, propositions.

Hinge propositions have their name from the metaphor that 'If I want the door [of our epistemic life] to turn, the hinges must stay put' (Wittgenstein, 1969, §343) just as a framework is fixed for a canvas to be stretched on it. That is, hinge propositions enable our rational epistemic life, our having beliefs, our investigating, and our giving reasons, like a hinge enables a door to turn. They are the presuppositions that we need to share in order to understand each other, or as Annalisa Coliva (2015) puts it:

they are constitutive of our rationality. Our hinges are the unmovable commitments that give meaning and doxastic force to our ordinary beliefs – they are the framework on which our belief system hangs. A further important notion is the idea of an *animal hinge* as proposed by Danièle Moyal-Sharrock (2004): animal hinges are the standing certainties which enable us to live our lives normally. For instance, the hinge proposition 'I have a body' is presupposed in almost everything that I do; I do not try to walk through walls, I simply pick up objects without even considering whether this is possible, etc. This shows how deep our hinges run and that they are mostly implicit.

What happens if our hinges nevertheless shift or diverge? If epistemic communities have divergent hinges, then they *deeply disagree* (Fogelin, 1985; Ranalli, 2020) and cannot rationally come to an accord. The divergent hinges mean that the disagreeing parties mean different things by what they are talking about: Take for example an atheist having the hinge that there is no god and a theist with the contradictory hinge arguing about miracles. They mean something altogether different by 'miracle': the former takes miracles to be definitionally impossible while the latter takes them to be normal occurrences. They would talk past each other about miracles.

If a single individual has a divergent hinge, and the hinge is sufficiently incompatible with the individual's community, then this looks a lot like a delusion as Campbell (2001, p. 96) suggests.[2] For the patient with Cotard, it is not an open question

[2] Bortolotti and Broome (2008, p. 835) suggest that the fact that a delusional person proffers different and incompatible reasons than the common-sense majority speaks against their delusion being a hinge. This presupposes a view of hinges as essentially socially anchored and shared institution to enable reason-giving. This would just be a confirmation of the rationalist thesis that the un-understandable delusion has replaced the ordinary common-sense hinge – doxastically, it functions like a hinge for the individual. I treat this concern which has also been raised by Thornton further below.

whether they are dead. Instead it is the hinge on which their other beliefs turn – just as it is a hinge for you that you are alive. This would colour the patient's experience of everything – given that you are dead, everything is meaningless. The shift in hinges also translates into a shift of the patient's language: 'My life is over' means something else if you have the hinge, i.e. presuppose, that *I am dead*, than if you presuppose that you are alive. Your hinges determine how you use certain terms and what inferential roles you assign them – for instance 'I am dead' does not entail 'I stopped breathing and moving' anymore. Thus, the patient's expression of the delusion appear un-understandable and bizarre. Campbell (2001, p. 98) compares this to a Kuhnian (1996) paradigm shift across which communication is not possible because the theoretical frameworks are incommensurable. Campbell consequently suggests that delusions may be divergent hinges that are produced top-down by an organic malfunction which have the far-reaching reverberations that manifest as delusional symptoms because they are hinges.

The key about this rationalist account of delusions is that it gives us an avenue to at least attempt to understand the delusional patient's belief system. If we find which hinges have shifted, we may at least try to model what the patient is believing and trying to say. Thus the benefit of the rationalist perspective is two-fold: First, it allows us to still ascribe a meaningful mental life to the delusional patient even though it may appear inscrutable to us. Second, it gives us a wedge to at least attempt to grasp what the delusional patient is trying to express. This is possible without having to take recourse to even more inscrutable strange experiences as the empiricist must, especially because it is often hard to see how a strange experience would

rationally support the specific content of a delusion over non-delusional alternatives like ‘I am hallucinating’ or ‘I am mentally ill’.

Criticisms of Rationalism

The most prominent and stringent critics of Campbell’s rationalism are Tim Bayne and Elisabeth Pacherie (2004). First they defend empiricism against Campbell’s criticism (Bayne and Pacherie, 2004, pp. 2–7). I will bracket this defence of empiricism here. Second, they directly criticise the tenability of Campbell’s theory.

Bayne and Pacherie (2004, p. 7) make an interesting and important observation about Campbell’s account: There are two logically separate aspects that we can examine each on their own merit. First, there is the idea that delusions are produced aetiologically top-down through a defectively formed belief that produces the further symptoms of delusion. Second, there is the epistemological idea that delusions are Wittgensteinian hinges which epistemically explain the delusion. Bayne and Pacherie criticise both these aspects one after the other. I agree that the two aspects should be distinguished; I suggest calling the former part of Campbell’s account his *aetiological top-down account*, while I would call the latter his *rationalist hinge account*. The account of delusion defended in (Campbell, 2001) is then the conjunction of the aetiological top-down account and the rationalist hinge account.

What are the objections against an aetiological top-down model of delusion? First, Bayne and Pacherie (2004, p. 8) raise the worry that it appears even less rational than an empiricist account. Namely, the empiricist bottom-up account gives the delusional patient at least their experience to point to as a reason for their delusions. But if the delusion is simply the product of a top-down cognitive defect, this is

completely arational and would not make the patient more understandable. Campbell can grant the point that the acquisition of the delusion is completely un-understandable because what he is aiming at is not the understandability of the acquisition of the delusion, but rather the interpretability of the solidified delusional belief system. To ascribe a delusional person a mental life, we do not need to understand how she became delusional, but we need to understand the delusional state. Also we ourselves sometimes discover that we just acquired a belief for no apparent reason – what is key for understandability is whether we maintain it also against defeaters.

The second kind of worry about the top-down account is about its aetiological nature (Bayne and Pacherie, 2004, p. 8): If delusional beliefs are the direct product of brain damage – why aren't there as many topics of delusions as there are topics of beliefs? And inversely, how could a supposed cognitive defect cause defects to non-cognitive autonomic systems as they occur in Capgras syndrome for instance? These aetiological worries are, I think, the weak point of the cognitive top-down account. (*see also* Hohwy, 2004) One first response is to note that there are more kinds of delusions with more diverse topics; especially delusional disorder (WHO, 2018, 6A24) can involve delusions about almost any topic. The second, a bit weak, response is to note that, differently from Cotard and Capgras, the hypothesised aetiologies for other kinds of delusions are much less clear.

Indeed, I think that limiting ourselves to Cotard and Capgras is not helpful when theorising about delusions as an epistemological phenomenon. Plausibly, there is no unified aetiology for our delusions: some may be top-down, others bottom-up, most will be mixed – depending on what caused the delusion. (Hohwy, 2004, pp. 67–68) It

would be surprising if drugs, schizophrenia, depression, neural damage, etc. created delusions all through the exactly same mechanism – be it top-down or bottom-up. Gerrans (2013, p. 87) indeed complains that the rationalist hinge part of Campbell's theory does not give us any neurocognitive explanation. Demanding such explanations, however, means misunderstanding the theoretical goals of epistemologists – in this the goal is to case integrate delusions with our epistemology and to explain how it relates to knowledge, doubt, certainty, and so on.

I want to argue that delusions are all unified by their epistemology – namely, all delusions are rationalist hinges. Consequently, I will defend the rationalist hinge part of Campbell's theory more stridently in this chapter. A broad complaint that Bayne and Pacherie (2004, p. 8) have about this rationalist thesis is that the Wittgensteinian hinge framework is not very well-developed – and indeed in 2001, hinge epistemology was in its infancy. Since then, a lot has happened, and hinge epistemology has come to be a field in its own right. Prominent contributions have been (Moyal-Sharrock, 2004; Coliva, 2015; Pritchard, 2016). These developments also allow us to better respond to the worries that Bayne and Pacherie raise about rationalism about hinges.

Bayne and Pacherie (2004, p. 8) agree that delusions have one key feature of hinge beliefs: They are extraordinarily resistant against counterevidence, and patients rather reject ordinary hinge beliefs than their delusion. For instance, someone with Cotard Syndrome and the hinge *I am dead* might reject the ordinary fundamental conviction that 'dead people do not breathe' given that they take themselves to be dead but breathing. This is exactly how hinges are supposed to work. It is not even

considered that the hinge certainty might be mistaken – instead the other beliefs are adapted.

However, Bayne and Pacherie (2004, p. 8) suggest that it is nevertheless implausible that delusions are hinges, because delusional patients frequently recognise the bizarreness of their delusion's content. Namely, if you ask them what they would think of someone else telling them a story analogous to their delusions with different protagonists they would frequently consider it to be bizarre and absurd, and they may admit that their delusion is hard to believe. Consequently, their delusion cannot be a hinge because that would arguably normalise the delusion's content for the patient.

The rationalist can respond that we need to pay careful attention to *what* the delusional hinge is. Note that delusions are often very subject-specific; they are *about* the patient; the patient is the delusion's protagonist. Consequently also the corresponding hinge is *about* the patient – the delusional content limitedly only applies to the patient. That is, a patient with persecutory delusion does not have the hinge that *people, including or like me, are being persecuted*, but specifically that *I am persecuted*. Similarly, Cotard Syndrome is not based on the hinge that *some dead people still move, speak, and breathe*. Instead, it has the hinge that *I am dead*, notwithstanding the fact that I still move, speak, and breathe. It is more epistemically conservative to take one's individual (hinge) case to be an exception – and to maintain that other dead people stay dead without breathing – conserving one's old hinges as much as possible. Consequently, the meaning of 'dead' would only undergo a patient-centric change.

This point relates to Bayne and Pacherie's (2004, p. 9) second worry: They point out that delusions are fairly encapsulated. Especially monothematic delusions like in Cotard and Capgras have often rather limited consequences for the patient's epistemic and practical life. The delusion may manifest mostly in linguistic behaviour and not engender many revisions in the patient's belief system. Thornton (2008, p. 162) makes this point especially vivid, and a telling illustration of this is the phenomenon of double bookkeeping (Fuchs, 2020, p. 76) where the patient maintains a normal epistemic and practical life but also has a delusional world-view in parallel which is kept insulated and separated from the former. If our delusions were hinges, then we would expect them to have far-reaching implications because they are our presuppositions about everything else.

I hinted already above that these delusional hinges are highly specific in their content: They are specifically only about the delusional individual, and they ascribe the individual a particular status, e.g. being dead. This specificity in content arguably blocks the spread of far-reaching epistemic and practical consequences. Consider for example the epistemic, semantic, and practical consequences of the delusion *I am chosen by god* – it is a hinge about me specifically and it does not imply that other people might also be chosen. My being chosen might even speak against other people's being so. Such encapsulated delusions are what Danièle Moyal-Sharrock (2004, p. 102) calls *personal hinges* which are mostly about the holder of the hinge – Wittgenstein's (1969, §486) example of a personal hinge is 'my name is L.W.' Note

also that as Bayne and Pacherie (2004, p. 6) themselves admit that also such specific delusions can have (horrific) practical consequences.[3]

The third worry that Bayne and Pacherie (2004, p. 9) raise is the semantic role of hinges. They criticise Campbell's suggestion that our delusions *qua* hinges have semantic consequences. If the delusional patient has hinges, e.g., *my spouse was replaced by an impostor*, then what 'impostor' and 'replaced' mean shifts in meaning. We cannot understand anymore what the patient means when talking about impostors. Replace 'impostor' with any delusional content you like. Intuitively however, the content of the delusion seems quite transparent: 'What does the Capgras patients believe if not that his wife has been replaced by an impostor?' (Bayne and Pacherie, 2004, p. 9) The threat is then that a rationalist hinge account exaggerates the semantic effects of a delusion.

I agree with Bayne and Pacherie, that delusions arguably do not have the far-reaching semantic consequences that Campbell appears to suggest. The delusional patient does not speak a profoundly different incommensurable language. Meanwhile, I do not think that such stark semantic consequences need to follow from rationalism about delusions. As argued above, the limited content of our delusions arguably also has limited semantic consequences. That is, a delusional patient does use some terms differently than the majority does, but this difference in use mostly manifests in differences of entailments which are limited to the consequences of the highly specific content of the delusion. What may occur in such a case is not a total

[3] In […], I argue for a weaker thesis that delusions are just psychologically indubitable certainties rather than hinge certainties, they consequently have weaker implications.

linguistic incommensurability, but rather local incommensurability. (Carey, 2009, p. 367)

A second, less prominent critique of Campbell's hinge rationalism is made by Thornton (2008). His argumentative target is, very limitedly, that a rationalist hinge account of delusions cannot provide us *any* epistemic access to understand a delusional person, thereby undermining the Davidsonian motivation for the rationalist hinge view. He motivates this with more broadly Wittgensteinian considerations about linguistic intelligibility, notably from the *Philosophical Investigations* (Wittgenstein, 1958). Note that the hinge rationalist is in no way wedded to the entirety of Wittgenstein's philosophy.

Thornton's (2008, p. 166) first point is that it is impossible to consider and understand false propositions as hinges because we do have our own set of hinges (note that delusional hinges need not necessarily be false). They would just be 'nonsensical sense', empty contradictions to our most fundamental certainties. I think this idea underestimates our capacities for modelling. For instance, our hinge framework is arguably Euclidian, and *space does not bend* is a hinge, nevertheless we are capable of doing non-Euclidian geometry. Bracketing our own hinges or limiting their reach just requires some imagination – we do it in mathematics and philosophy (Coliva and Doulas, 2022), why should we not be able to do it in psychiatry?

Thornton's (2008, p. 170) second argument is analogously that we are incapable of conceiving of delusions as abnormal framework propositions because this would be incommensurable with our linguistic practice. He considers two options: First, that delusions generate a totally alien and globally incommensurable hinge framework.

He rejects this by appealing to Davidson and McDowell without elaboration; I think we can reject the global option because some communication with delusional patients is always possible.

Second, he considers the option – that I appealed to above – that delusional hinges are very local and narrow. Thornton (2008, p. 177) argues that hinges work just like grammatical rules (*cf.* Coliva, 2015). His key idea appears to be that grammatical rules essentially need to be socially shared in order to function as grammatical rules – if they are not then the grammar breaks down and there are no rules left at all. Consequently, hinges also need to be shared in order to function as hinges, otherwise they are no hinges. Given that delusions are definitionally not shared, they cannot be hinges that could be understood.

To this, the rationalist about delusions can respond that Thornton confuses philosophy of language with epistemology. Social epistemic practice, i.e. the sharing and transmission of knowledge, obviously is very central to epistemology, but it is not the case that epistemic activity completely breaks down if sharing breaks down. Language's constitutive function is the sharing of information; meanwhile there is a legitimate individual epistemology – internalism is not dead yet. In that sense, Thornton's analogy between grammatical rules of language and epistemological hinge rules falls short. While the linguistic function of grammatical rules is completely undermined by not being shared, the function of epistemic hinge rules is only damaged and reduced if it is not shared, because their social epistemic function is undermined. Nevertheless, delusions that form hinges which are not shared are simply dysfunctional but not completely undermined.

To illustrate this point: If you are very, very smart (I am not) then you could play chess against yourself just in your head – and thereby follow the rules of chess. You would be playing chess, even though Wittgenstein and Thornton would complain that it is meaningless to talk about such an internal game of chess – but talking about it is different from playing the game itself.

Still, the worry remains that locally not shared hinges would completely undermine the understandability, i.e. social epistemological aspects, of the delusional hinge belief system. The ordinary shareability of beliefs is undermined, a delusional patient's testimony fails given the discrepancies in the speaker's and hearer's respective frameworks. Nevertheless, I can reiterate my point above: Our imagination is not that limited – just as we can suspend or model alternative hinge frameworks in philosophy or geometry, we can do so in psychiatry in order to salvage some meaning from the patient's testimony. (see also Henriksen, 2013)

As third and final objection, Thornton (2008, pp. 172–173) argues against Eilan's (2000) rationalist hinge account, which I will present below, that we cannot be considered to believe our hinges.[4] Many authors agree with Thornton that hinges cannot be believed because they are not reasons-sensitive or knowledge-apt. (Coliva, 2015, p. 44; Pritchard, 2016, p. 92) Note however, that also delusions in this sense should not be considered to be beliefs – they are also not reasons-sensitive or knowledge-apt. The question whether we believe hinges and delusions is vexed (see Chapter 20), but it appears likely that if you fall on one side of the question for hinges, then you will fall on the same side for delusions and *vice versa*.

[4] Similarly, Henriksen (2013, p. 112) suggests that we only assume hinges – but assumptions are easily defeated by counter-evidence, so hinges cannot be assumptions.

In sum, we can split Campbell's (2001) view into two parts: the aetiological top-down account about what produces or causes delusions, and the rationalist account about the epistemic role that delusions play in a patient's belief system. While some cases of delusion may be produced top-down, the bottom-up role of experience cannot be denied.[5] Consequently, the top-down account of delusions does not apply generally. However, I have suggested that the rationalist hinge account of delusions applies to all cases of delusions, and on this view delusions are unified by being pathological hinge certainties. As you will see below, this rationalist hinge account has been quite attractive to theorists in the last two decades.

This hinge rationalism about delusions has two principal theoretical consequences: First, we can make predictions about how delusions, *qua* hinges, relate to other beliefs. Namely our ordinary beliefs should be interpreted in the light of the delusion, ordinary beliefs should be rejected or relativised if they contradict a delusion, and the delusion is simply presupposed and taken for granted, not requiring any argument. Second, if delusions are hinges, then beliefs that may appear to be delusions but which do not play the functional role of hinges do not count as full-blown delusions. Instead they would simply be irrational, and maybe pathological, beliefs.

Developments of Rationalism

The first development of rationalism about delusions appeared roughly at the same time as Campbell's (2001). Namely, Naomi Eilan (2000) examines the epistemic status of schizophrenic delusions and the problem of radical interpretation that this

[5] I personally do not think that a pure empiricist account can get off the ground because we arguably also want to distinguish delusions from ordinary mistaken beliefs based on hallucinations. Consequently, some top-down mechanism must play a role.

raises. Her starting point is a puzzle that Karl Jaspers' account of delusions generates because Jaspers emphasises the un-understandability or utter bizarreness of schizophrenic delusions while at the same time wanting to ascribe a mental life to delusional patients.[6] Eilan (2000, pp. 106–107) considers the empiricist route but argues that not all schizophrenic delusions could be boiled down to strange and unsettling experiences. Eilan therefore looks to Campbell's work, going with rationalism.

She points out that someone with divergent hinges would appear to be mad, because they have such an alien take on the world. (Eilan, 2000, p. 103) However they would not become completely un-understandable. Namely, Eilan emphasises that while delusions as hinges explain *prima facie* why they are un-understandable, the hinge account also allows us to 'fall in, to an extent, with a deluded subject's reasoning' (Eilan, 2000, p. 109) by imagining a delusion's content to be a hinge certainty.

Further, Eilan (2000, p. 112) suggests that what makes a particular strange belief into a delusional hinge is an emotional loading of the proposition with significance. For instance in the case of morbid jealousy (Kingham and Gordon, 2004), the belief that *my partner cheats on me gets* loaded by the emotion of jealousy which transforms the testable and refutable belief into an incontrovertible hinge certainty which colours all my further beliefs and experiences. Consequently, Eilan does not endorse the top-down account of delusion because she also looks to emotion as a

[6] This somewhat simplifies Jasper's account, who distinguishes different possible kinds of understanding.

causal factor, nevertheless she is a rationalist because she takes delusions to be hinges.

Another prominent early development of Campbell's rationalism was Klee's (2004) suggestion that delusions split into *pedestrian* and *stark* delusions. Pedestrian delusions are cases that are tractable and do not fall out of our shared epistemic framework, e.g. litigious delusions that *the state has been treating me unjustly*. Klee (2004, p. 29) explains them as instances of a Davidsonian fragmented belief set where the different fragments do not interact epistemically. Meanwhile, stark delusions are the extreme cases that are un-understandable, e.g. claims that *external forces insert thoughts into my mind* or that *I am dead*. These, Klee suggests, are instances of the patient having acquired a different set of hinges from us – thereby making starkly delusional patients un-understandable.

Hohwy (2004, pp. 66–67) criticises that the line between stark and pedestrian delusions is much harder to draw than Klee suggests. Notably, the distinction seems to cut across particular types of delusions. For instance, some paranoid delusions are very pedestrian – *the NSA is wiretapping my phone calls* – while others are completely out of this world and unexpected – *the Spanish Inquisition is monitoring my every thought*. I agree with Hohwy; additionally, I think that even clearly pedestrian delusions can be explained by delusional hinges with a very narrow content. Consequently, a pure hinge rationalism is more economical than Klee's mixed view. In this vein, Bardina (2018) develops an account of the broad range of possible kinds of delusions of varying starkness by relying on Moyal-Sharrock's (2004) typology of different kinds of hinges – personal hinges, animal hinges, etc.

A prominent alternative account of the role of hinges for delusions is by Rhodes and Gipps (2008, 2011). Instead of describing delusions as hinge certainties like Campbell, they emphasise the delusional patient's loss of the ordinary hinges that we all share. Thus, a delusion is a lack of shared hinges, which permits the delusional patient to have such strange beliefs as 'I am setting the sun'. An interesting recent elaboration on this idea is by Jeppsson (2021) who reports that she experienced her own delusional episodes as the loss of her hinge bedrock. She also points out that endorsing scepticism involves a suspension of our fundamental hinge certainties – e.g. 'there is an external world' or 'there are other minds' – and suffering from a delusion feels, according to her, a lot like being in the throes of scepticism.

Bortolotti (2011, p. 83) objects to Rhodes and Gipps's proposal that it does not explain the fixedness of our delusions. If we lacked fundamental hinge certainties nailing our beliefs down, should our delusions not become florid and constantly change?

An indirect response to this objection can be found in Fuchs's (2020). He rejects the rationalist label because he misinterprets Campbell's account as involving irrational thinking errors and because he rejects the top-down model. (Fuchs, 2020, p. 71) However, he endorses the rationalist hinge framework, integrating it into an enactivist two-factor account. Fuchs suggests that schizophrenic delusions develop in two stages: First, as proposed by Rhodes and Gipps (2008) and Jeppsson (2021) our hinge certainties become inoperative, and our perception is not structured by our animal hinges anymore. (Fuchs, 2020, p. 67) However, this is not yet the full-blown delusion, but a mere precursor. Second, the patient attempts to reorder their

unhinged epistemic life by settling for new hinge certainties – the full blown delusional hinge which integrates the patient's unsettled experience again. (Fuchs, 2020, p. 68)

Fuchs has an additional reason why he rejects the 'rationalist' label for his hinge account. Namely, he argues similarly to Coliva (2015), that abandoning our ordinary hinges, which are constitutive of rationality, means abandoning rationality entirely. Note that Bortolotti (2005) argues *contra* Campbell's motivation for rationalism and empiricism that we can ascribe delusional patients beliefs even though the patient is patently irrational. Coliva's view of rationality, however, is relying on a very narrow and fundamental range of hinges which guarantee classical logic and a minimal common-sense world view. I believe that many delusions remain at least partially within the bounds of this minimal framework.

Finally, in […], I also develop an epistemological rationalist account of delusions. Notably, the paper gives a positive epistemological argument that delusions are really (hinge) certainties. It shows that dysfunctional hinge certainties are the best available epistemic candidate for delusions by examining and rejecting the proposed alternative accounts of the epistemology of delusions. A key idea is that delusions are taken to be dysfunctional hinges – that is delusions pervert the function of hinge certainties just as autoimmune diseases pervert the function of our immune system. As a weaker option, it also suggests that delusions are just certainties that the patient cannot doubt, and not hinge certainties. Meanwhile, the paper remains silent on the aetiological roots of delusions and does not defend a top-down account; instead defending a rationalist hinge theory of delusions.

Conclusion

Campbell's rationalism has proven to be very influential. Principally, it has served as a foil for its opponents – especially discussion of the top-down/bottom-up distinction has proven very fruitful. Meanwhile, also the rationalist hinge aspect of Campbell's account has played an influential role. The basic Wittgensteinian idea that delusions are defective hinges has seen many interesting developments which leaves hinge rationalism about delusions as a distinct epistemological position in the philosophical debate about delusions.

References

Bardina, S. (2018) 'Abnormal Certainty: Examining the Epistemological Status of Delusional Beliefs', *International Journal of Philosophical Studies*. Routledge, 26(4), pp. 546–560. doi: 10.1080/09672559.2018.1497072.

Bayne, T. and Pacherie, E. (2004) 'Bottom-Up or Top-Down: Campbell's Rationalist Account of Monothematic Delusions', *Philosophy, Psychiatry, & Psychology*, 11(1), pp. 1–11.

Bortolotti, L. (2005) 'Delusions and the Background of Rationality', *Mind and Language*, 20(2), pp. 189–208. doi: 10.1111/j.0268-1064.2005.00282.x.

Bortolotti, L. (2011) 'Continuing Commentary: Shaking the Bedrock', *Philosophy, Psychiatry, & Psychology*, 18(1), pp. 77–87.

Bortolotti, L. and Broome, M. R. (2008) 'Delusional Beliefs and Reason Giving', *Philosophical Psychology*, 21(6), pp. 821–841. doi: 10.1080/09515080802516212.

Campbell, J. (2001) 'Rationality, Meaning, and the Analysis of Delusion', *Philosophy,*

Psychiatry, & Psychology, 8(2), pp. 89–100.

Carey, S. (2009) *The Origin of Concepts*. Oxford: Oxford University Press.

Coliva, A. (2015) *Extended Rationality*. Basingstoke: Palgrave Macmillan.

Coliva, A. and Doulas, L. (2022) 'What philosophical disagreement and philosophical skepticism hinge on', *Synthese*. Springer Netherlands, 200(3), p. 251. doi: 10.1007/s11229-022-03735-6.

Davidson, D. (1968) 'On saying that', *Synthese*, 19(1/2), pp. 130–146.

Eilan, N. (2000) 'On understanding schizophrenia', in Zahavi, D. (ed.) *Exploring the Self. Philosophical and psychopathological perspectives on self-experience.* Amsterdam: John Benjamins Publishing Company, pp. 97–113. doi: 10.1075/aicr.23.09eil.

Fogelin, R. (1985) 'The Logic of Deep Disagreements', *Informal Logic*, 7(1), pp. 1–8.

Fuchs, T. (2020) 'Delusion, reality and intersubjectivity: A phenomenological and enactive analysis', *Phenomenology and Mind*, 18(1), p. 120. doi: 10.17454/pam-1810.

Gerrans, P. (2013) 'Delusional Attitudes and Default Thinking', *Mind & Language*, 28(1), pp. 83–102. doi: 10.1111/mila.12010.

Gipps, R. G. T. and Rhodes, J. (2008) 'The Background Theory of Delusion and Existential Phenomenology', *Philosophy, Psychiatry, & Psychology*, 15(4), pp. 321–326.

Gipps, R. G. T. and Rhodes, J. (2011) 'Delusions and the Non-epistemic Foundations of Belief', *Philosophy, Psychiatry, & Psychology*, 18(1), pp. 89–97.

Henriksen, M. G. (2013) ‘On incomprehensibility in schizophrenia’, *Phenomenology and the Cognitive Sciences*, 12(1), pp. 105–129. doi: 10.1007/s11097-010-9194-7.

Hohwy, J. (2004) ‘Top-Down and Bottom-Up in Delusion Formation’, *Philosophy, Psychiatry, & Psychology; Psychology*, 11(1), pp. 65–70. doi: 10.1353/ppp.2004.0043.

Jeppsson, S. (2021) ‘Psychosis and intelligibility’, *Philosophy, Psychiatry and Psychology*, 28(3), pp. 233–249. doi: 10.1353/ppp.2021.0036.

Kingham, M. and Gordon, H. (2004) ‘Aspects of morbid jealousy’, *Advances in Psychiatric Treatment*, 10(3), pp. 207–215.

Klee, R. (2004) ‘Why Some Delusions Are Necessarily Inexplicable Beliefs’, *Philosophy, Psychiatry, & Psychology*, 11(1), pp. 25–34.

Kuhn, T. (1996) *The Structure of Scientific Revolutions*. 3rd edn. Chicago: The University of Chicago Press.

Moyal-Sharrock, D. (2004) *Understanding Wittgenstein’s On Certainty*. Basingstoke: Palgrave Macmillan.

Pritchard, D. (2016) *Epistemic Angst*. Princeton: Princeton University Press.

Quine, W. V. O. (1960) *Word and Object*. Cambridge, Massachusetts: MIT Press.

Ranalli, C. (2020) ‘Deep disagreement and hinge epistemology’, *Synthese*. Springer Netherlands, 197(11), pp. 4975–5007. doi: 10.1007/s11229-018-01956-2.

Stone, T. and Young, A. W. (1997) ‘Delusions and brain injury: The philosophy and psychology of belief’, *Mind and Language*, 12(3–4), pp. 327–364. doi: 10.1111/j.1468-0017.1997.tb00077.x.

Thornton, T. (2008) ‘Why the idea of framework propositions cannot contribute to an understanding of delusions’, *Phenomenology and the Cognitive Sciences*, 7(2), pp. 159–175. doi: 10.1007/s11097-007-9079-6.

WHO (ed.) (2018) *ICD-11 for Mortality and Morbidity Statistics (ICD-11 MMS)*. Available at: https://icd.who.int/browse11/l-m/en.

Wittgenstein, L. (1958) *Philosophical Investigations*. Oxford: Blackwell.

Wittgenstein, L. (1969) *Über Gewissheit - On Certainty*. Edited by G. E. M. Anscombe and G. H. von Wright. Oxford: Blackwell.