

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/165460>

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Understanding Delusions: Evidence, Reason, and Experience

by

Chenwei Nie

A thesis submitted in partial fulfilment of the requirements
for the degree of
Doctor of Philosophy

Department of Philosophy

University of Warwick

September 2021

Contents

Acknowledgements	iii
Declaration	vi
Abstract	vii
Chapter 1. The challenges of delusions	1
1. Introduction	1
2. The evidence challenge.....	3
3. The specificity challenge	11
4. Some clarifications	15
Chapter 2. A critique of current theories of delusions	21
1. Introduction	21
2. The anomalous experience hypothesis.....	25
3. The reasoning abnormality hypothesis	61
4. Something is missing	85
Chapter 3. Believing what we clearly perceive	91
1. Introduction	91
2. Phenomenal clarity compels assent: first pass.....	97
3. Phenomenal clarity is a distinctive phenomenon.....	100
4. How does phenomenal clarity compel assent? An inquiry	115
5. Proposing a dual-force framework of beliefs	133
6. The dual-force account of akratic beliefs.....	136
Chapter 4. The dual-force framework for explaining delusions	152
1. Introduction	152
2. Personal remarks on delusions	152
3. The dual-force account of delusions.....	157
4. Phenomenal clarity, salience, and dopamine.....	176
5. Conclusion	182
Appendix. Are delusions beliefs?	186
1. Introduction	186
2. The puzzling features of delusions.....	188
3. The deadlock between doxasticists and anti-doxasticists	191
4. Several attempts to break the deadlock	196
5. What next?	199
6. Thinking of delusions as seeming-based beliefs	203
Bibliography	208

Acknowledgements

I am very lucky and enormously grateful to be supervised by Johannes Roessler. The supervision meetings with him have been the best part of my PhD life. He has always been encouraging and supportive when I explored some novel ideas; but remained quite candid and critical when I sometimes made myself slip into unpromising ones. At the writing up stage of this Thesis, I have received tremendous help from him. He has read many drafts of each chapter and patiently guided me on how to improve the Thesis: from the overall structure to the use of the semicolon. For a long time, we often met once per week. I only wish one day I could become such a devoted and inspiring supervisor as he is.

My PhD project has also benefited greatly from the regular meetings with my Graduate Progress Committee, including Thomas Crowther, Guy Longworth, Christoph Hoerl, and Stephen Butterfill. I am particularly grateful to Naomi Eilan for her invaluable guidance during Term 2, 2018, when I was anxiously preparing for the MPhil to PhD upgrade.

I first learned about delusions from Anne Aimola Davies and Martin Davies during the China-UK-US Philosophy Summer School in 2013, and later continued my study under their guidance when I was a recognised student at Oxford from 2015 to 2016. I owe them a deep debt of gratitude. In the past years, I have also gratefully received mentorship from Nick Bunnin, John Hyman, Huiming Ren and Jie Liu; and have been very fortunate to have met and received generous help from many brilliant philosophers, including John Preston, Max De Gaynesford, Luke Elson, Nadine Elzein, Maria Alvarez, Matthew Parrott, Max Coltheart, Huaping Wang, Liwu Rong, Wenyu Xie, Yongjun Fu, Qixiang Huang, Jie Tian, Shanbo Wang, Zhangyin Li,

Peng Guo, Peter Su, Tongli Wu, Daniel Lim, Kristjan Laasik, Jiwei Ci, Alexander Bird, Lisa Bortolotti, Ryan McKay, Fabienne Peter, Peter Poellner, Matthew Nudds, Diarmuid Costello, Hemdat Lerman, David Bather Woods, and Lucy Campbell. Special thanks go to Kirk Surgener, Barnaby Walker, Quassim Cassam, Christoph Hoerl, Stephen Butterfill, Adina Covaci, and Johannes Roessler, for helping me learn about how to teach undergraduate seminars.

Parts of the work in the Thesis have been presented in Shandong University, East China Normal University, King's College London, University of Warwick, University of Essex, Friedrich Schiller Universität Jena, and the 27th and 28th Annual Meetings of the European Society for Philosophy and Psychology. I thank the audience for the very helpful discussions.

At last, I would like to thank the Doctoral College and the Department of Philosophy at the University of Warwick for the generous Chancellor's International Scholarship; thank the Lanchester Library at the University of Coventry, where I wrote some early parts of this Thesis. Many thanks to all my colleagues, flatmates, and friends, in particular Dafni Zirou, Emin Gok, Beatrice Pagliarone, Peaw Khemchira Kaitwatcharachai, Mert Yirmibes, Giulia Luvisotto, Ahilleas Rokni, Giulia Lorenzi, Chris Noonan, Daniel Mario Weger, Daniele Cassaghi, Antonio Danese, Sanja Sreckovic, Matt Tompkins, Seren Rimmington, Gemma Louise Benevento, Guangyu Xi, Michelle Yuanbo Liu, Hongjie Zhu, Huizhong Sun, Wenli Yao, Ke Xia, Junlong Liu, Rong Wang, Ruiyun Zeng, Ye Ping, Xiaoqi Wen, Jinyang Zhou, and Pengchao Gao for their kindness and company; to Rachel Harrington, Clare Simpson, Emily Hargreaves, Victoria Cox, Sarah Taylor, and Kerry Law for the unswerving administrative support; to my parents, Haisheng Nie and Huanyun Li, and my sister, Chenrong Nie, for their love; and most of all to Jie Lu for being my sunshine on rainy days and being my star during dark nights.

This thesis is dedicated to all who suffer from delusions.

Declaration

I, Chenwei Nie, confirm that the thesis is my own work. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. The thesis has not been submitted for a degree at another university.

Abstract

This thesis develops a novel framework for explaining delusions.

In Chapter 1, I introduce the two fundamental challenges posed by delusions: the evidence challenge lies in explaining the flagrant ways delusions flout evidence; and the specificity challenge lies in explaining the fact that patients' delusions are often about a few specific themes, and patients rarely have a wide range of delusional or odd beliefs.

In Chapter 2, I discuss the strengths and weaknesses of current theories of delusions, which typically appeal to one or both of two factors: anomalous experience and reasoning abnormality. I argue that anomalous experience can help explain the specificity of delusions, but has difficulties in addressing the evidence challenge; reasoning abnormality can help address the evidence challenge, but has difficulties in explaining the specificity of delusions. This suggests that there may be an important factor that has not been captured by current theories of delusions.

To search for this missing factor, in Chapter 3, I turn to normal believing. Inspired by the literature on Cartesian clarity and phenomenal dogmatism, I develop a dual-force framework of believing, according to which beliefs can be understood as the results of the interaction between the justificatory force and causal force of evidence and the justificatory force and causal force of clear experience, in which something clearly seems to be so to the subject. This framework suggests that the missing factor may be the clear experience with its distinctive phenomenal clarity that compels assent.

In Chapter 4, I return to delusions, and argue that the dual-force framework can help us to get a better grip on some personal descriptions of delusions; make progress in addressing the evidence and specificity challenges of delusions; and shed new light on the underpinnings of delusions. In the end, I conclude with some remaining questions for future study.

Chapter 1.

The challenges of delusions

The most profound distinction in psychic life seems to be that between what is meaningful and allows empathy and what in its particular way ununderstandable, 'mad' in the literal sense.

(Jaspers, 1913/1997, p. 577)

They [i.e. patients with a psychotic condition] make me angry and I find myself irritated to experience them so distant from myself and from all that is human.

(Freud, 1928, in a letter to Istvan Hollos, quoted in Dupont, 1988, p. 251)

1. Introduction

Mental disorders both fascinate and baffle inquisitive human minds, especially those of philosophers and psychiatrists. On the one hand, mental disorders offer rare opportunities to unveil the nature of the human mind; on the other hand, some mental disorders appear so baffling that it seems to be a formidable task for one to have a grasp on what is going on in them. The bafflement is famously reflected in Jaspers' remark that many mental disorders are ununderstandable (Jaspers, 1913/1997; see also Davies & Coltheart, 2000; Eilan, 2000; Hoerl, 2019; Walker, 1991). Among those (seemingly) ununderstandable mental disorders, the paradigmatic, and arguably most baffling, are delusions, which "[s]ince time immemorial," writes Jaspers, "ha[ve] been taken as the basic characteristic of madness" (Jaspers, 1913/1997, p. 93; see also Berrios, 1991).

Delusions are notoriously held by the subjects with a peculiar sense of finality. As Coltheart et al. (2011) write:

The delusional patient espouses her belief with a sense of absolute knowing, as if the delusional content could not be other, and so is beyond any requirements for objective justification. The delusional belief appears to be experienced as self-evident, without the need for any justification of its truth, and the patient seems unable to even entertain the possibility that the belief might be incorrect (p. 281)

This peculiar finality is most strikingly manifested in the flagrant ways delusions flout the norms concerning the relationship between beliefs and evidence, or the flagrant ways delusions flout evidence for short. The latter clinically defines delusions: according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), “[d]elusions are fixed beliefs that are not amenable to change in light of conflicting evidence.” (American Psychiatric Association, 2013, p. 87).

To explain delusions, in large part, is to explain where this peculiar finality comes from, and particularly to explain why delusions flout evidence.¹ The rest of this chapter will focus on getting clear what the main challenges of delusions are; and the rest of the thesis will discuss to what extent current theories of delusions can address these challenges (Chapter 2), and develop a new framework which may help make some progress (Chapters 3 and 4).

¹ This is, of course, not the only bafflement revolving around delusions. There are many other puzzles concerning the relationship between delusions and actions, and between delusions and the subjects’ other mental states, such as their emotions, desires, and other beliefs. For example, a patient, who had the delusion that the doctors and nurses were trying to poison her, “may happily consume the food they give her” (Sass, 1994, p. 21). As regards the patient’s action, it is puzzling why the patient did not reject the food given she believed that the food may be poisonous; as regards the patient’s emotion, it is puzzling why the patient happily ate the food. Addressing these puzzles is, however, not the focus of this thesis. I will briefly discuss some of them in Appendix.

2. The evidence challenge

Delusions flout evidence in many baffling ways. Let us start by considering the following examples:

- Capgras delusion: “This woman [the subject’s wife] is an imposter” (Capgras & Reboul-Lachaux, 1923; translated in Ellis et al., 1994; for reviews, see Edelstyn & Oyebode, 1999; Pandis et al., 2019);
- Cotard delusion: “I’m dead” (Cotard, 1882; for reviews, see Debruyne et al., 2009; Young & Leafhead, 1996);
- Delusion of thought insertion: “Alien thoughts, such as ‘Kill God’, are put into my mind” (Frith, 1992; Roessler, 2013; for reviews, see López-Silva, 2018; Mullins & Spence, 2003);
- Delusion of supernumerary phantom limb: “I have three arms” (Halligan et al., 1993; for a recent review, see Kim et al., 2017).

These delusions are conventionally called bizarre delusions (for reviews, see Cermolacce et al., 2010; Spitzer et al., 1993; Yuksel et al., 2018). When we lay our eyes on them, these delusions immediately strike us as bizarre. The reasons why they strike us as bizarre may, to some extent, vary from case to case: it is highly unlikely that one’s wife is an imposter; it is “pragmatically self-defeating” to believe that oneself is dead (Bermudez, 2001, p. 479); it is far-fetched and technically impossible that thoughts can be inserted into one’s mind; it is biologically improbable that a person has three arms. Nonetheless, one of the common reasons may be what has been called the “sheer implausibility” of the contents of bizarre delusions (Coltheart et al., 2011, p. 281).

To say that the contents of bizarre delusions are utterly implausible is, however, not to say that they are absolutely impossible

by any notions of possibility. The ideas of someone being an imposter, the living dead, thoughts being inserted, and having three arms are all familiar themes in movies and scientific fictions; some of them were received as real not so long ago in human history; and perhaps some are still held as possible in certain cultures or places (see e.g. Reimer, 2009). The idea of someone being an imposter, for example, has been traced back to Greek mythology, in which the most infamous of all may be Zeus' ability of impersonation in his unrelenting pursuit of love affairs (Christodoulou, 1986; see also Oyebode, 2021, p. 1). For someone who believed in those stories, it was certainly not absolutely impossible for a person to be an imposter.

The contents of bizarre delusions are utterly implausible in the sense that they flout common knowledge which is possessed by nearly everyone in the society (Vanderschraaf & Sillari, 2021). Bizarre delusions are what Parrott (2016) calls *epistemic impossibilities*, as opposed to epistemic possibilities which “are those things that are possible given what is known or, equivalently, those things that are compatible with what is known ... [by the] subject's epistemic community.” (pp. 282-285) When we lay our eyes on the contents of bizarre delusions, we immediately know that they are utterly implausible in the real world because they flout common knowledge, according to which such bizarre contents are only possible in, say, imaginary worlds (Dubourg & Baumard, 2021).

Common knowledge is arguably a basic source of evidence. When someone posts on his Facebook that he has successfully designed a perpetual motion machine, and only needs one million pounds to make it, we do not need to give him one million pounds to make a machine so that we can have some evidence, e.g. a machine that fails to produce perpetual motion, to disconfirm his claim; rather, our knowledge that a perpetual motion machine is impossible is good evidence disconfirming his claim. What is so bizarre in bizarre delusions is the fact that bizarre delusions are formed and maintained

in spite of the fact that they flout the good evidence provided by common knowledge. In this sense, to hold a bizarre delusion is like to believe that the claim about the perpetual motion machine is true. (And to act on one's bizarre delusion is like to actually give the charlatan one million pounds to make the imaginary perpetual motion machine.)²

In contrast to bizarre delusions, some delusions may have relatively mundane contents that do not flout common knowledge at first sight. Consider the following delusions:

- Delusional jealousy: “My partner is unfaithful to me” (Todd & Dewhurst, 1955; for reviews, see Enoch et al., 2020; Qureshi et al., 2016);
- Persecutory delusion: “My neighbours are plotting against me” (for reviews, see Freeman & Garety, 2014; Murphy et al., 2018).

One's partner being unfaithful and one's neighbours being plotting against them are certainly (and unfortunately) not unfamiliar phenomena, and are compatible with common knowledge. What distinguishes a delusion with some relatively mundane content from a non-delusional belief with the same content is the fact that people with delusions have little, if any, evidence for their delusions and face significant counterevidence. The belief that “My partner is unfaithful” is a normal belief if the subject has good evidence, such as the partner's confession or a trustworthy detective's testimony, that it is the case; it is a delusion if the subject has no such kind of evidence, and all the subject's evidence suggests that the partner is faithful.

Measured by their compatibility with common knowledge, the difference between bizarre delusions and more mundane delusions

² A belief's flouting of common knowledge is not a sufficient condition for it to be a bizarre delusion. I do not claim that when a person believes that a perpetual motion machine is possible, they suffer from a delusion.

arguably comes in degrees, and they should be located towards the opposite ends of the spectrum. Measured by their relationship with evidence, however, both appear to be close at the same end of the spectrum.

Delusions' flouting of evidence should not be confused with cases in which the subject avoids any evidence that may contradict their beliefs. When a scholar refuses to read any critique of their views, this kind of action is frowned upon, but they may not be deluded in holding that their views are true. When an authoritarian government asks its officials to censor everything that may defy what it claims to be true, this kind of action is deplorable, but the authoritarian government may not be deluded in believing that what it claims to be true is true (nor are the people deluded, in the clinical sense, in believing that what the government claims to be true is true).³ By contrast, many people with delusions, in various baffling ways, do engage with the evidence.⁴ Let us consider these baffling ways in detail.

First, many people with delusions may respond to the request for evidence, even though the "evidence" provided by them is mostly too weak or elusive to support their delusions. For example, some people with the Capgras delusion may claim that the imposter exhibited certain distinguishing features which "can be seen by certain details... a little mark on the ear... a thinner face... a longer moustache... different colour eyes... the way of speaking... the way of walking" (Ellis et al., 1994, p. 129); and when they were confronted with the invalidity or elusiveness of these claims, some may add that "you have got to be quick to notice" the distinguishing features of the

³ The term *delude* has a common meaning that is synonymous with *deceive*. In this sense, the people are deluded or deceived by the authoritarian government. In this thesis, however, I only use the terms *delusion*, *delude*, *delusional* in their clinical senses.

⁴ In this sense, not only is the authoritarian government morally blameworthy for its action, but the way it treats evidence makes it epistemically more blameworthy than people suffering from delusions.

imposter (Young et al., 1993, p. 696). It is baffling how this kind of “evidence” could convince a person that his wife is an imposter.⁵

Certainly, not all people with delusions would respond to the request for evidence. But it seems to me that the percentage of those who would give some response might be more sizable than what has been presented in the literature. For example, Parnas (2004) argues that patients with what he calls the Autistic-Solipsistic Delusion not only rarely engage in convincing others that their delusions are true, but “may even be puzzled by the need for any justification.” (p. 158) While I think Parnas is right about the existence of the Autistic-Solipsistic Delusion, it is contestable whether the patient in the exemplary case he uses to illustrate the Autistic-Solipsistic Delusion in fact responded to the request for evidence. Parnas chooses the following case reported by Janet (1929):

Madeleine, a patient described in detail by Janet, had a habit of walking on her tiptoes, as a part of her experience of ‘divine ascension’. Janet once provokingly commented that if her ‘divine ascension’ was really true, then her feet should be at least 10 centimeters above the floor. Madeleine responded: ‘What a strange idea, applying metric measures to divine matters!’ (pp. 146-147)

In his comment on the case, Parnas (2004) writes: “It is clear that although the patient seems to claim certain state of affairs, she does not make these claims in order to seek intersubjective agreement.” (p. 157) It seems to me, however, that there is a sense in which, although Madeleine’s reason might not be a good reason, by offering reason to address Janet’s challenge, Madeleine was actually

⁵ An alternative way to understand these claims about the distinguishing features of the imposter is that they are not the evidence based on which the Capgras delusion is formed, but the subjects’ confabulations that are posterior to the formation of the Capgras delusion. For the relationship between delusions and confabulations, see e.g. Turner and Coltheart (2010) and Coltheart (2017).

trying to convince Janet that her delusion was true. In this sense, Madeleine did not seem to suffer from an Autistic-Solipsistic Delusion. Instead, Madeleine could be taken as a patient who responded to the request for evidence.

Second, many people with delusions may appear to have *partial* insight into the force of the conflicting evidence. This is taken by some scholars as an essential characteristic of monothematic delusions. For example, Corlett and Fletcher (2021) write:

[P]atients with monothematic delusions appear to understand what sort of evidence might undermine their beliefs but do not modify them in the face of such evidence. They do not appear to take on and use such disconfirmatory evidence. This, of course, is one of the key criteria for establishing that a delusion is present so presumably, *if the patients did not show this effect, they would not be considered to be deluded.* (p. 233, emphasis added)

This baffling phenomenon is evident in the following case in which the person FE suffered from the delusion of mirrored-self misidentification: when he looked at the mirror, he believed that he saw a different person rather than his own reflection.

FE's family tried on numerous occasions to dissuade him from his belief by providing him with evidence contrary to the delusion. FE would listen attentively to their arguments and often agreed with their logic, but his delusional belief remained steadfast. (Breen et al., 2000b, pp. 82-83)

Third, many people with delusions may even appear to have *partial* insight into the implausibility of delusions (David, 1990; Jolley & Garety, 2004; Startup, 1997). Regarding people with schizophrenia, Startup (1997) reported that they had “little difficulty” in identifying

other people's delusions as symptoms of mental illness, even though they were unable to apply the same standards to their own delusions (p. 203). Similarly, Feyaerts et al. (2021b) reported that, in their interview, the participants with schizophrenia "were often well aware that delusional experience would be judged as bizarre or unlikely when set against normal evidential standards." (pp. 794-795) Regarding people with the Capgras delusion, Young (1998) reported: "If you ask 'What would you think if I told you that my wife has been replaced by an impostor?,' you will *often* get answers to the effect that it would be unbelievable, absurd, an indication that you had gone mad." (p. 37, emphasis added) In a similar vein, Halligan et al. (1993) reported that when questioned about the patient's delusion of supernumerary phantom limb, the patient's replies included that "It sounds nonsense I know." (p. 163)⁶

Fourth, in some cases, delusions may act as the subjects' reason to dismiss the counterevidence. For example, Young and Leafhead (1996) reported that when they suggested to the patient JK, who had the Cotard delusion, that the fact that she had feelings "surely represented evidence that she was not dead, but alive. JK said that since she had such feelings even though she was dead, they clearly did not represent evidence that she was alive." (p. 158)

In summary, delusions' flouting of evidence is manifested in many baffling ways, including the facts that delusions lack supporting evidence, and face significant counterevidence, which may include common knowledge in the case of bizarre delusions; delusions may remain steadfast even when the subjects appear to have *partial* insights into the force of the counterevidence and the implausibility of

⁶ It is unclear whether there are cases in which the patients are not only aware of the implausibility of their delusions from the perspective of others, but also aware that their delusions are in fact false. In such cases, the patients may say something to the effect that "I know *my* delusional belief is false, but I believe it." To my knowledge, no such cases are reported. But, see Krstić (2019), for an argument that the patient Mr. F's grandiose delusions, reported by Joseph et al. (2011), might be interpreted as such.

the contents of their delusions; delusions may even act as the subjects' reason to dismiss the counterevidence. These flagrant ways delusions flout evidence constitute a main challenge to theories of delusions, which I will call *the evidence challenge*.

Admittedly, when a particular case is concerned, what counts as a piece of evidence itself can be contestable. For example, so far I have assumed the fact that the person looks exactly like your wife counts as a piece of evidence that she is your wife. By contrast, in the literature on delusions, it is widely assumed that the fact that this woman looks exactly like your wife is a piece of evidence that confirms both the wife hypothesis and the imposter hypothesis because “[a]fter all, a *good* impostor would look like the patient’s wife” (Aimola Davies & Davies, 2009, p. 269, emphasis added; see also Davies & Egan, 2013, pp. 701-702). In addition, patients with the Capgras delusion are found to have reduced affective responses to the “imposter” (Ellis & Young, 1990; see also Brighetti et al., 2007; Ellis et al., 1997; Ellis et al., 2000; Hirstein & Ramachandran, 1997; Nuara et al., 2020). And the reduced affective responses are often taken as a piece of evidence that confirms the imposter hypothesis but disconfirms the wife hypothesis.

However, we may wonder whether the reduced affective responses could be taken as a piece of evidence that disconfirms both the wife hypothesis and the imposter hypothesis because, like the patient’s wife, an *excellent* impostor would induce normal affective responses in the patient too.

Following this analysis, it seems that whether the wife’s look and the patient’s reduced affective responses would count as the evidence that confirms or disconfirms the imposter hypothesis is dependent on the patient’s conception of the impostor, particularly on whether the patient thinks the impostor is an ordinary impostor who would not be able to look exactly like his wife, a good impostor who would look like his wife but would not induce normal affective

responses in him, or an excellent imposter who would both look like his wife and induce normal affective responses in him. Consequently, it seems likely that what counts as a piece of evidence may vary from case to case, and may turn out to be less straightforward than we would hope even when a particular case is concerned.

Nonetheless, I think the controversies revolving around a particular piece of evidence would not change the key point that it is a general characteristic of delusions that they flout evidence in many flagrant ways. For example, even in a particular case of the Capgras delusion, the patient thinks that the “imposter” is really good at impersonating his wife such that his delusion is not disconfirmed by the fact that the “imposter” looks exactly like his wife, it remains true that there is ample evidence that his delusion flouts in many baffling ways: the Capgras delusion is implausible in the light of common knowledge; it is at odds with the fact that the “imposter” knows things that only his wife knows; it is also at odds with the testimonies from his friends and relatives; and the patient may even have partial insights into the force of the counterevidence and the implausibility of the delusion. This general characteristic of delusions will be one of the primary concerns of this thesis; and I will set aside the complex issues concerning a particular piece of evidence.

3. The specificity challenge

Delusions’ flouting of evidence might make one wonder whether the subjects simply lose the general ability to properly evaluate beliefs in the light of evidence. This, however, does not seem to be the case because if the subjects lost such a general ability, they would lack lots of normal beliefs and end up with a wide range of odd or delusional beliefs. The latter, however, does not happen to people with delusions (Gold & Hohwy, 2000). As a matter of fact, delusions are usually restricted to a few specific themes, which are well-

recognised by psychiatrists (e.g. Bentall et al., 1991; Frith, 1992), although there are different ways we can classify these themes.

Cutting (2003), for example, divides characteristic schizophrenic delusions into eleven themes: (1) delusional perception; (2) thought withdrawal; (3) thought insertion; (4) thought broadcasting; (5) made feelings; (6) made impulses; (7) made volitional acts; (8) somatic passivity; (9) bizarre delusions; (10) multiple delusions; (11) widespread delusions. (p. 18, Table 2.4)

Recently, Gold and Gold (2015) propose to classify delusions into twelve themes: (1) persecutory delusions; (2) delusional jealousy; (3) erotomanic delusions; (4) religious delusions; (5) grandiosity; (6) delusions of control; (7) delusions of thought (including the delusion of thought insertion); (8) somatic delusions; (9) nihilistic delusion (including Cotard delusion); (10) delusions of guilt or sin; (11) delusions of reference; and (12) misidentification delusions (including Capgras delusion) (pp. 115-117; see also Gold & Gold, 2012; Stompe et al., 2003).^{7, 8}

A dramatic form of the specificity of delusions is manifested in monothematic delusions (as opposed to polythematic delusions), in which case the subject only has one delusion or a few delusions about the same theme (Coltheart et al., 2007; Coltheart, 2013; Davies et al.,

⁷ Gold and Gold (2015) further argue that these themes are primarily *social* themes. In line with this social view, some have proposed that certain abnormalities in social processes, such as testimonial abnormalities, are the main contributory factor in the aetiology of delusions (Bentall et al., 1991; Bell et al., 2021; Miyazono & Salice, 2020; Williams, 2020; Williams & Montagnese, preprint). I am not committed to the social views. For critique, see e.g. Coltheart and Davies (2021b, pp. 10-11), Reed et al. (2020), and Corlett (2021).

⁸ When discussing delusions that follow brain injury, Stone and Young (1997) put forward a different classification, according to which delusions involve four basic themes of human beliefs: “the person’s body (e.g. thinking that your arm is someone else’s), the environment (e.g. thinking that you are somewhere other than where everyone around you claims to be), the self (e.g. thinking that you are dead), or other people (e.g. thinking that your wife has been replaced by an impostor).” (p. 327) It is, however, worth noting that delusions’ involving these four basic themes does not mean that there are delusions of every imaginable kind. This is because each of the four basic themes contains a wide range of sub-themes, and delusions appear to be limited to only a few of the sub-themes.

2001; Radden, 2011, Chapter 2), and the subject “seems otherwise perfectly normal” (Coltheart et al., 2011, p. 282), and “perfectly in touch with reality” (Stone & Young, 1997, p. 329). As Davies and Egan (2013) put it, “monothematic delusions [are] islands of delusion in a sea of apparent normality” (p. 690). The booming of the study of monothematic delusions starts with the development of cognitive neuropsychiatry (David & Halligan, 1996; Ellis & Young, 1990; Halligan & Marshall, 1996). This phenomenon is, however, arguably recognised in history. For example, in *An essay concerning human understanding*, Locke (1690/1975) wrote: “[I]t comes to pass, that a Man, who is very sober, and of a right Understanding in all other things, may in one particular be as frantick, as any in *Bedlam*” (p. 161, emphasis in original).

The specificity of delusions raises a variety of questions. For example, we may wonder why delusions are limited to specific themes, and why there are no “delusions of every imaginable kind” (Gold & Hohwy, 2000, p. 157); we may wonder why delusions are limited to these themes, such as the twelve themes proposed by Gold and Gold (2015), rather than other themes; we may also wonder why an individual suffers from, say, delusions about three of Gold and Gold’s twelve themes, but has no delusions about the rest of the twelve themes; and we may wonder why some people suffer from monothematic delusions while others suffer from polythematic delusions.

In this thesis, I do not intend to address these questions. Instead, I am particularly concerned with the question about the discrepancy between the way people with delusions respond to the evidence related to their delusions and the way they respond to the evidence related to their non-delusional beliefs. Put it another way, my main concern is the question of why people with delusions, on the one hand, flagrantly flout the evidence related to their delusions, but, on the other hand, appear normal as regards their non-delusional

beliefs. I call the challenge to answer this question *the specificity challenge*. As we have seen, this challenge is posed by both polythematic delusions in schizophrenia⁹ and monothematic delusions.¹⁰ Regarding monothematic delusions, it has also been called “the monothematicity problem” (Davies et al., 2001, pp. 149-154).

In view of the specificity challenge, the evidence challenge becomes even more formidable, because the specificity of delusions suggests that people with delusions do not suffer from a general disability in evaluating beliefs in the light of evidence, and their flouting of evidence is somehow restricted to their delusions.

The flouting of evidence and the specificity constitute the two fundamental challenges to theories of delusions. The rest of the thesis will focus on discussing to what extent current theories of delusions can address these challenges, and to what extent what I will call the dual-force framework may help make some progress. Before we move to that, there are several other issues concerning the nature of delusions which are worthy of some clarifications.

⁹ Some may seem to have assumed that polythematic delusions in schizophrenia do not pose a form of the specificity challenge. For example, while accepting that it is problematic to think of a domain-general cognitive bias as the main factor in the aetiology of monothematic delusions because monothematic delusions are limited to specific themes, Noordhof and Sullivan-Bissett (2021) write: “This is not a problem for those suffering from polythematic delusions as a result of schizophrenia, and so it is no surprise that bias theories were first formulated in this context.” It is not clear to me what their reason is. I think that, when a subject with schizophrenia has a few delusions about, say, five different themes, it still poses a challenge to the view that these delusions are caused by a domain-general cognitive bias.

¹⁰ One might wonder whether polythematic delusions pose a less baffling form of the specificity challenge. The answer is, however, not straightforward. This is because it seems unclear which of the following cases is more baffling: a person with one delusion about one single theme, or a person with a few delusions about five themes.

4. Some clarifications

4.1. Are delusions really beliefs?

The aforementioned bafflement of delusions may make one wonder whether delusions are really beliefs. If delusions are not beliefs, then it opens up many possible ways to explain why people with delusions flout evidence. If delusions are, say, imaginings, then we may not be so baffled by the fact that they are not based on evidence.

Indeed, in the philosophical literature, there is an ongoing debate on the doxastic nature of delusions (for reviews, see Bortolotti & Miyazono, 2015; Bortolotti, 2010). Apart from the flagrant ways delusions flout evidence and the specificity of delusions, the debate also revolves around other puzzling features of delusions such as the incongruences between patients' delusions and their intentions, affective states, and behaviours. For example, many patients with the Capgras delusion did not go out to search for their real spouses, did not seem to have the intention to do so, and sometimes the patient even "had a friendly and warm relationship with" the imposter (Dietl et al., 2003, p. 462). All these puzzling features of delusions raise doubts about whether delusions are beliefs.

I think that delusions are beliefs. But my arguments differ from those in the literature. In Appendix, I will present a critical analysis of the current debate, and I will argue that the debate has encountered a persistent deadlock. In the end, with the help of the philosophical tools that are developed in Chapter 3 and Chapter 4, I will argue that thinking of delusions as seeming/clear experience-based beliefs can help us make some progress in explaining the puzzling features of delusions. Before that, it is important to note that, in the scientific

literature, delusions have been largely assumed to be beliefs.¹¹ I shall follow this assumption.

4.2. The heterogeneity of delusions

There is a wide variety of delusions. We have mentioned bizarre delusions, delusions with more mundane contents, monothematic delusions, and polythematic delusions. In addition, we may find people with elaborated delusions or delusional systems, or people with circumscribed delusions in which case there is no obvious connection or interaction between the delusions and the subjects' other beliefs. Monothematic delusions are often circumscribed; polythematic delusions are often elaborated. But, as emphasised by Davies et al. (2001), the notions of circumscription and elaboration are different from the notions of monothematicity and polythematicity; and they point out that there are empirical cases in which the person has “a monothematic but somewhat elaborated delusional system” (p. 136).¹²

Delusions can occur in a number of medical conditions, including, but not limited to, schizophrenia, dementia (Cipriani et al., 2014), epilepsy (Elliott et al., 2009b, 2009a), Parkinson's disease (Factor et al., 2014), traumatic brain injury (David & Prince, 2005), and stroke (Torrise et al., 2019). The Capgras delusion, for instance, has been reported in all these medical conditions and more (for reviews, see Bell et al., 2017; Currell et al., 2019; Darby & Prasad, 2016; Josephs, 2007; Pandis et al., 2019)

¹¹ Some cognitive neuropsychological theories do offer brief philosophical discussions about whether delusions are beliefs (see e.g. Davies & Coltheart, 2000; Davies et al., 2001; Stone & Young, 1997).

¹² It is not clear whether there are cases in which the person has polythematic but relatively circumscribed delusions.

Delusions may be induced by drugs (Corlett et al., 2010a; Voce et al., 2019), and may also be transiently induced by hypnosis (Coltheart et al., 2018; Connors, 2015).

The heterogeneity of delusions might give one the impression that “both supporting and contradictory evidence for almost any theory can be found among them.” (Adams et al., 2021, p. 2) While this is an exaggeration, it is true that the vast differences among delusions raise the question of what strategies one should adopt to study delusions. “[A] useful strategy,” write Young and Leafhead (1996), “may ... be to try to understand in detail a particular, relatively tightly defined phenomenon, and then see whether one's explanation can be broadened to encompass other observations.” (p. 150; see also Young, 2000)

In the literature, two-factor theories (Aimola Davies & Davies, 2009; Coltheart, 2007; Coltheart et al., 2011; Davies, 2009; Davies & Coltheart, 2000; Davies et al., 2001; Davies & Egan, 2013; Langdon & Coltheart, 2000; McKay, 2012; Nie, 2016, 2019) have been focusing on monothematic delusions, particularly those “of neuropsychological origin” (Davies & Egan, 2013, p. 690; for the application to polythematic delusions in schizophrenia, see Coltheart, 2013; Coltheart et al., 2007); the predictive processing theories (Corlett et al., 2010b; Corlett et al., 2016; Sterzer et al., 2018) have been focusing on delusions in schizophrenia; some theories have been focusing on a particular theme of delusions: for example, what can be called cognitive-bias theories have been focusing on persecutory delusions, which are the most common delusions in schizophrenia (Freeman & Garety, 2014; for the application to delusions in general, see Ward & Garety, 2019).

Despite the heterogeneity of delusions and the apparent discrepancies among the focuses of these theories, it is crucial to note that the flouting of evidence and the specificity are two shared characteristics of delusions. Hence it is essential for any theory of

delusions to explain these shared characteristics. In this thesis, I shall evaluate a theory of delusions by discussing how well it can address the evidence challenge and the specificity challenge, and by and large set aside the issues related to the factors that aim to explain the heterogeneity of delusions.

I shall focus on the monothematic, circumscribed, bizarre Capgras delusion, for two considerations. First, Capgras delusion is a widely discussed delusion in the literature. It is what Bayne (2017) calls “the poster child of contemporary cognitive neuropsychiatry” (p. 82). Second, the most formidable form of the evidence challenge is posed by bizarre delusions, as opposed to more mundane delusions; the most formidable form of the specificity challenge is posed by monothematic and circumscribed delusions, as opposed to polythematic and elaborated delusions; the most challenging kind of delusions is, therefore, monothematic, circumscribed, bizarre delusions. In this sense, the monothematic, circumscribed, bizarre Capgras delusion is an apt choice to test how well a theory of delusions can address the evidence challenge and the specificity challenge.

4.3. Other variations

There are other variations, even when we focus on a particular kind of delusions with the help of the aforementioned taxonomy. First, the conviction of delusional beliefs may come in degrees, which can be measured by psychometric examination (e.g. Combs et al., 2006). For example, Green et al. (2018) recently asked a group of 17 subjects with delusions, who had committed crimes related to their delusions, to rate their delusions according to a scale (0% = strongly disbelieve, 25% = disbelieve, 50% = neither believe nor disbelieve, 75% = believe, and 100% = strongly believe). The result was: 8 subjects

reported 100%; 1 reported 95%; 2 reported 90%; 3 reported 75%; 1 reported 60%; and 2 reported 0% (p. 322).

Second, the flouting of evidence may vary in degrees. There are reported cases in which the subjects with delusions appear to doubt or even abandon their delusions in the face of counterevidence. For example, Coltheart (2007) reported an interview with a patient who believed that his wife M. was a stranger looking like his wife. During the interview, the patient said: "The lady knows me way back. She couldn't say things that happened 40 years ago, and I wonder where she gets them from. And then I worked it out and I've wondered if it's M. all the time. It's nobody else." (p. 1054)

Third, delusions may wax and wane. Coltheart (2007) reported a case in which the patient "sometimes expressed the belief that his mother, father, and sister had been replaced by impostors, but at other times correctly identified them as genuine family members." (p. 1053; for a similar case, see e.g. Mackie et al., 1994; see also David, 1999; Sharp et al., 1996)

The degree of delusional conviction, the degree of the flouting of evidence, and the wax and wane of delusions may co-vary: a patient may hold their delusions with a low degree of conviction, may flout the evidence to a low degree, and may eventually reject the delusions. But it is worth noting that these three phenomena are, strictly speaking, independent of each other: a patient with a high (or low) degree of conviction of their delusions may flout the evidence to either a low or high degrees, and vice versa; and delusions may wax and wane without obvious reason.

Will these variations be a problem for theories focusing on paradigmatic cases in which the evidence is flouted to a high degree, and the delusions are held with strong convictions and do not frequently wax and wane? I do not think they would be a grave problem. The main reason is that such variations in delusional cases, in principle, can be explained by the variations in the causal factors.

A lesser degree of delusional conviction, for example, might be explained by a mild form of the proposed cognitive deficits that cause the conviction (see e.g. Coltheart, 2007; Coltheart et al., 2011).

To sum up, in this chapter I have introduced the two fundamental challenges of delusions: the evidence challenge lies in explaining the flagrant ways delusions flout evidence; the specificity challenge lies in explaining the fact that patients' delusions are often about a few specific themes, and they rarely have a wide range of delusional or odd beliefs. I have also made some clarifications about further complications concerning the doxastic nature, the heterogeneity, and some variations of delusions. I suggest that these complications do not seem to be obstacles for a theory to start by focusing on explaining the paradigmatic cases of delusions. The primary concern of this thesis will be the question of what the causal factors in the paradigmatic cases of delusions are, rather than how the proposed factors can further accommodate the complications. As I will argue in the next chapter, we still have not arrived at a satisfactory answer to the former question.

Chapter 2.

A critique of current theories of delusions

[I]ntuitive notions of folk psychological intelligibility do struggle somewhat with the idea of a belief that is brutally caused by disease or injury. We can stay close to the intuitive conception of beliefs as being adopted on the basis of experience and inference if the element of brute causation enters one step earlier and disease or injury causes abnormal experiences. We can also avoid a struggle with intuitive notions of folk psychological intelligibility if disease or injury perturbs reasoning processes in ways that are not too far removed from the familiar effects of mood, tiredness or alcohol, for example.

(Davies & Coltheart, 2000, p. 8)

1. Introduction

Theories of delusions typically appeal to one or both of two factors to explain delusions: anomalous experience and reasoning abnormality.¹³ The discussion of the potential roles of anomalous experiences in the aetiology of delusions can be dated back at least to Jaspers. He (1913/1997) wrote:

The environment [from the perspective of some patients] is somehow different—not to a gross degree—perception is unaltered in itself but there is some change which envelops everything with a subtle, pervasive and strangely uncertain light. ...This

¹³ I use the term *reasoning* in a broad sense including both personal-level conscious reasoning processes, such as the kind of reasoning we employ when we do mathematics and philosophy, and subpersonal-level unconscious reasoning processes, such as Bayesian inference, which, some argue, is the way information is processed in our brain (see e.g. Friston, 2012).

general delusional atmosphere with all its vagueness of content must be unbearable. Patients obviously suffer terribly under it, and to reach some definite idea [delusion] at last is like being relieved of some enormous burden. (p. 98)

Here, the anomalous experience is the delusional atmosphere whose vague content inflicts an enormous burden on the patients. And delusions with definite contents arise to relieve the enormous burden caused by the delusional atmosphere with the vague content.

We may wonder whether it is true that the determining factor at play is the vagueness or definiteness of the content. But the more general idea is that anomalous experiences do harm to the patients whereas delusions bring benefits to the patients.

Jaspers, however, cautioned against exaggerating the explanatory power of this analysis of delusions in terms of the benefits. The first reason offered by Jaspers was that “[i]t is doubtful whether the foregoing analysis will hold in all cases.” (p. 98) By “all cases”, Jaspers might have in mind variants of delusions in schizophrenia, in some of which the delusional atmosphere is not present.¹⁴ His basic point, nonetheless, has been further corroborated by the abundance of reported cases of monothematic delusions since the late 20th century, which are often not associated with the delusional atmosphere (see e.g. Breen et al., 2000b).¹⁵

The second reason was that despite the benefit of burden relief, it is hard to find an intelligible link between a patient’s anomalous

¹⁴ For a recent account of how the delusional atmosphere in schizophrenia may lead to the crystallization of delusions, see Henriksen and Parnas (2018).

¹⁵ There are some reported cases in which the monothematic delusions were associated with certain anomalous experiences whose phenomenal characters may overlap with the phenomenal characters of the delusional atmosphere. Stone and Young (1997), for example, argue that the anomalous experience of the subjects with the Capgras delusion is associated with some suspicious mood, which is a key phenomenal character of many current notions of the delusional atmosphere (Henriksen & Parnas, 2018).

experience and delusion (see also Davies & Coltheart, 2000; Eilan, 2000; Hoerl, 2019; Walker, 1991). When we believe something for a reason, there is usually an intelligible link between the belief and the reason. Such an intelligible link is, however, absent in the relationship between the anomalous experience and delusion. The vague content of the anomalous experience does not seem to qualify as an intelligible reason for holding a delusion. Nor is an intelligible link present in the relationship between the benefit of burden relief and the delusion: for one thing, the benefit of believing that *p* arguably does not qualify as a (good) reason for believing that *p*; for another, even if we accept that the benefit of believing that *p* can be a reason for believing that *p* (see e.g. Pascal, 1670/1995; Rinard, 2018, 2019), it is still far from clear why the patient favours the delusions over abundant non-delusional beliefs, such as that the patient has mental disorders, which may explain the anomalous experience and relieve the associated burden too.

The third reason was that delusions also do serious harm to patients' well-being, and overall delusions bring about more harm than benefits.¹⁶ For example, according to recent reviews (Karakasi et al., 2019; Pandis et al., 2019), the Capgras delusion in schizophrenia spectrum disorder is associated with terror and shame, and with a moderate to high level of violence including attempted murder, for which the subjects faced serious (legal) consequences.¹⁷

¹⁶ Bortolotti (2015, 2016, 2020) argues that in certain cases delusions can deliver some epistemic benefit. This, however, does not change the main point that “[d]elusions ... are generally harmful, impairing good functioning and causing anxiety and distress.” (Bortolotti et al., 2016, p. 48)

¹⁷ Jaspers' idea about delusions and their possible benefits appeared to be shared by many psychiatrists of his time. For example, Southard (1912) wrote: “A delusion as we see it might assert itself as a belief by becoming useful, consistent with evolution, pragmatic.” (p. 328) This idea enjoys renewed attention from the recent discussion concerning whether delusions might bring about psychological, social, and biological benefits, which in turn might be contributory factors in the aetiology of delusions (Bortolotti, 2016, 2020; Ritunnano et al., 2021; Westermann et al., 2018). And these benefits are also taken to be one of the main reasons for the view that delusions are (evolutionarily) adaptive (Clamor & Krkovic, 2018; Fineberg & Corlett, 2016; Gunn & Bortolotti, 2018; Lancellotta & Bortolotti, 2019; McKay & Dennett, 2009; Mishara & Corlett, 2010). I acknowledge that there are cases in which a

The discussion of the potential roles of reasoning abnormalities, broadly defined, in the aetiology of delusions can be dated back at least to Kraepelin (1889). At first sight, it may have some intuitive appeal that certain reasoning abnormalities must play an important role in the aetiology of delusions, because delusions appear to be something that one with proper reasoning ability would reject. It turns out to be extremely difficult, however, to articulate exactly what the reasoning abnormalities are. A pioneering contribution was made by Von Domarus (1944), who proposed that people with schizophrenia, of which delusion is a main symptom, suffer from a deficit in syllogistic reasoning. Apart from syllogistic reasoning, later studies also explored possible abnormalities in other forms of logical reasoning, including propositional/conditional reasoning and probabilistic reasoning. Somewhat surprisingly, however, a growing number of logical reasoning experiments showed that people with schizophrenia, by and large, did not perform worse than ordinary people in logical reasoning. In a review published in 2000, Mujica-Parodi et al. argue that “research to date has been inconclusive on the fundamental question of whether patients with delusional ideation show abnormalities in logical reasoning.” (p. 73; see also Maher, 2001) In an updated review published in 2014, Gangemi and Cardella hold a similar view that “the specific [logical reasoning] deficits in schizophrenia remain to be shown”. (p. 107)

Despite the problems with these early proposals, the general idea that anomalous experiences and/or reasoning abnormalities play important roles in the aetiology of delusions is retained in current theories.¹⁸ In the next two sections, I shall discuss to what extent

particular delusion might be associated with certain benefits. But, I think, Jaspers’ insightful critique is still applicable to the modern and more sophisticated versions of analysing delusions in terms of the benefits.

¹⁸ Another once-dominant approach to delusions is the psychodynamic approach which is rarely discussed since the late 20th century. For critical reviews, see Stone and Young (1997), de Pauw (1994), and Ellis and de Pauw (1994).

current notions of anomalous experience and reasoning abnormality can help us explain delusions. In brief, I shall argue that anomalous experiences can help address the specificity challenge but have difficulties in addressing the evidence challenge; reasoning abnormalities can help us address the evidence challenge but have difficulties in addressing the specificity challenge. The two-factor theory takes into account both anomalous experiences and reasoning abnormalities and makes significant improvements, but it also inherits some of the shortcomings of both.¹⁹ This suggests that there must be something in delusions, which has not been captured by the notion of anomalous experience and the notion of reasoning abnormality. In the last section, I shall discuss some of the proposed candidates for the missing factor. My own proposal will be developed in Chapters 3 and 4.

2. The anomalous experience hypothesis

2.1. Maher's one-factor theory of delusions

To discuss the modern frameworks for explaining delusions, a starting point is Maher's (1974, 1988, 1992, 1999) one-factor theory, according to which delusions arise as *normal* explanations of *anomalous* experiences.

Compared to the view discussed by Jaspers that some delusions arise to relieve the enormous burden caused by the delusional atmosphere, thinking of delusions as explanations is quite

¹⁹ In the literature, it is sometimes assumed that the proposed factors in current theories of delusions can sufficiently explain delusions. For example, Flores (2021) holds that it is because of the "continued strange experiences, motivational factors, and cognitive biases, ... the delusion stays in place, or does not sufficiently change, in the face of counterevidence." (see also van Loon, 2021) By contrast, I think it is a misleading exaggeration to think of delusions as a kind of beliefs that are entirely based on the proposed factors. In this chapter, I will argue that the proposed factors fall short of sufficiently explaining delusions.

a distinctive view.²⁰ This view is not without precursors. Coltheart et al. (2010) traced it back to James' (1890/1950) *The Principles of Psychology*, in which James wrote: "The delusions of the insane are apt to affect certain typical forms, very difficult to explain. But in many cases they are certainly *theories* which the patients invent to account for their bodily sensations" (chap. XIX, emphasis added).

Maher's one-factor theory, however, does not merely reinstate the view that delusions are the subjects' theories to explain their anomalous experiences. According to Maher (1988), delusions are a particular kind of theories: "normal theories" or normal explanations (p. 20). What does Maher mean by normal theories or normal explanations? He (1974) writes:

It is the core of the explanations (i.e. the delusions) of the patient are derived by cognitive activity that is essentially indistinguishable from that employed by non-patients, by scientists, and by people generally. . . . [A] delusion is a hypothesis designed to explain unusual perceptual phenomena and developed through the operation of normal cognitive processes. (p. 103)

A variety of cognitive processes may be employed when we explain something: such as attention, remembering, reasoning, and metacognition (Horne et al., 2019). Among them, Maher emphasises that the reasoning processes in the aetiology of delusions are normal. To highlight this point, I will call it the normal-reasoning view (instead of the normal-cognition view).

²⁰ Maher (2006) seems to accept that certain benefits of having a delusion are another factor in the aetiology of delusions. He writes: "The delusional explanation provides enough relief from anxiety that it becomes difficult for the individual to abandon it and return to the initial state of confusion and distress. Thus, strong resistance develops to counter-evidence that could rebut the delusional belief." (p. 182)

The normal-reasoning view was also held by some other psychiatrists around the early twentieth century. Southard (1912), for example, wrote the following sentences which are strikingly similar to Maher's:

Delusion-formation ... is a reaction not far removed from normal reactions, a perversion of the believing process, distinguished from that of the true believers (i.e., ourselves, the sane public and proper constituents of society) in little. (pp. 327-328)

Distant kin of the normal-reasoning view were arguably prominent in the 17th and 18th centuries, during which period many believed that "delusions resulted from failures (caused by physical reasons) in the apparatus that served to acquire experiences. The system designed to process information was in order." (Berrios, 1991, p. 7) For example, in *An Essay Concerning Human Understanding*, Locke (1690/1975) wrote:

[M]ad Men ... do not appear to me to have lost the Faculty of Reasoning: but having joined together some Ideas very wrongly, they mistake them for Truths; and they err as Men do, that argue right from wrong Principles. (p. 161)

For Maher, the normal-reasoning view is not merely a view but was supported by the repeated failure of the empirical studies to find logical reasoning abnormalities in schizophrenia (Maher, 2001; Mujica-Parodi et al., 2000). This normal-reasoning view comes with a firm prediction: that is, if an ordinary person has a similar anomalous experience, they would develop some similar delusions.

To examine this prediction, it is crucial to get clear what Maher means by anomalous experience. There are four important points. First, Maher's anomalous experience is a *conscious* experience: patients with delusions are consciously aware of the anomaly

associated with their experience. Second, the anomalous experience is *not* anomalous in the sense that it provides the patients with distorted shapes, colour, sounds, and the like. This is in line with Jaspers' (1913/1997) observation that "perception itself remains normal and unchanged" in patients with delusions (p. 100). Third, the anomaly is an anomalous feeling of significance. According to Maher (1999), feelings of significance are a distinctive kind of primary experiences, in the sense that "they have the same quality of irreducible directness as do such experiences as sensory experience of color, the feeling of physical pain, the experience of sound, and other sensations." (p. 552) At the subconscious level, feelings of significance are generated by a monitoring mechanism when it detects "discrepancies between expected and perceived input experience" (p. 556). Fourth, the feeling of significance is imprecise. It is "a *vague general feeling* that prompts us to look for [an explanation]." (p. 554, emphasis in original)

To illustrate how feelings of significance give rise to beliefs in non-delusional cases, Maher offers the following example:

A common experience that we encounter occurs when we meet somebody familiar to us, but have the conscious experience that the person "looks different somehow." We scan the person's appearance and may fail to detect any concrete differen[ce]. We ask, and find out that our friend has changed hair-style, shaved off a beard, grown a beard, or so forth. Sometimes our scan has already identified the difference and we do not need to ask. This writer once shaved off his beard and the next day was greeted with comments ranging from direct recognition of what had changed, to the case of one colleague who commented "There's something different about you. Oh yes. You are wearing a different kind of tie." (p. 553)

This is a helpful example to compare with the Capgras delusion: the belief that this woman looks like his wife but is not his wife. According to Maher, both the ordinary belief in the quoted case and the Capgras delusion arise as *normal* explanations of the experiences with feelings of significance. The difference between them is that, while the feeling of significance in the quoted case is a normal feeling, the feeling of significance in the Capgras delusion is something that normally would (and should) not have occurred, and hence is anomalous.

Despite the fact that the comparison between the quoted non-delusional cases and the Capgras delusion helps us have a grasp on what Maher thinks is going on in delusions, it does not suffice to tell us, however, what exactly the anomalous experience is such that the Capgras delusion arises as a normal explanation of it. The feeling of significance_d in the Capgras delusion is unlikely to be the same as the feeling of significance_n in the quoted non-delusional case. By definition, the feeling of significance_n is prevalent in our daily lives. If the feeling of significance_d were the same as the feeling of significance_n and the Capgras delusion were a normal explanation of the feeling of significance_n, then the belief that someone is an imposter would have been a familiar belief in our daily lives. The latter is, of course, not true. Therefore, the anomalous experience in the Capgras delusion must be a distinctive kind. Now the question is: how can we know what exactly this distinctive kind of anomalous experience in the Capgras delusion is?

2.2. Searching for the anomalous experience in the Capgras delusion

To find out what a person's experience is, one natural way is to ask the person. But people with delusions are often reluctant or find it difficult to give detailed and accurate reports of their

experiences: the reason why they are reluctant to report may include the worry about the social stigma of mental illness; and the reason why they find it difficult to report may include the fact that their experiences are too anomalous to be described by the ordinary language, which, one might argue, has not evolved to describe something so alien to people without delusions. Sometimes when patients are relatively engaging in an interview, their reports, such as the claim of the patients with the Capgras delusion that they noticed a little mark on the imposter's ear, sound more like confabulations rather than the actual anomalous experiences that give rise to their delusions.^{21, 22}

Because of these obstacles and many others, in the past psychiatrists and philosophers had to heavily rely on intuitions and conjectures, which unfortunately often conflicted with each other. With little consensus on how they can be empirically verified or disconfirmed, these theories often ended up in chaos. Similar chaos is infamously evident in the literature on psychodynamic theories of mental disorders. In their review of psychodynamic theories of the Capgras delusion, Ellis and de Pauw (1994) complained: "Over the last seven decades, a plethora of mutually-incompatible psychodynamic explanations, reflecting the imagination and conceptual frameworks of individual authors, have been invoked to

²¹ For a review of the difficulties in finding out, via interview, what the anomalous experiences in schizophrenia are, and some ways that may help overcome the difficulties, see Pienkos et al. (2021).

²² Even ordinary people's self-reports of ordinary experiences are arguably unreliable (Schwitzgebel, 2008). For example, in an experiment Moore and Schwitzgebel (2018) examined people's reports of their reading experiences. They found that "there is substantial variability in reports about reading experience, both within and between participants", and the reports varied from "visual imagery experiences, [to] inner speech experiences, and to experiences of conscious visual perception of the words on the page" (p. 57). One interpretation of this finding is that people have diverse reading experiences. In contrast, the alternative interpretation is that people have the same reading experience, and the diverse reports are due to the unreliability of our ability to give accurate self-reports of our experiences. If the latter interpretation is true, it would add another dimension of complications to the task of finding out, via interview, what the anomalous experiences in delusions are.

account for the CD [Capgras delusion].” (p. 318; see also de Pauw, 1994; Koritar & Steiner, 1988; Todd, 1982)

The major breakthrough is made by Ellis and Young (1990), with the help of which we now have a better understanding of what the anomalous experience in the Capgras delusion may be. Ellis and Young propose that the Capgras delusion is the “mirror image of prosopagnosia” with covert recognition (p. 244; see also Young & Ellis, 1989).

People with prosopagnosia are unable to overtly recognise familiar faces, such as the faces of famous people, their friends, family members, or even their own (Ellis, 1989).²³ However, some people with prosopagnosia may exhibit forms of covert recognition of familiar faces (for reviews, see e.g. Bruyer, 1991; Rivolta et al., 2013; Schweinberger & Burton, 2003): for example, when they see familiar (versus unfamiliar) faces, some retain a higher level of activity in the autonomic nervous system, which is measured by skin conductance (see e.g. Bauer, 1984; Bobes et al., 2004; Jones & Tranel, 2001; Tranel & Damasio, 1985; Tranel & Damasio, 1988).²⁴

According to Ellis and Young’s (1990) “mirror image” proposal, in comparison to the individuals with prosopagnosia who are unable to overtly recognise familiar faces but retain normal covert autonomic responses to familiar faces, people with the Capgras delusion can overtly recognise familiar faces but have abnormal covert autonomic responses to familiar faces (p. 244).²⁵ While it is true that many people

²³ This is an impairment restricted to the domain of face recognition. Patients with prosopagnosia do not lose the general ability to recognise people’s identities. Typically they can use voices or other features, such as clothes, to help them identify people.

²⁴ Not all people with prosopagnosia exhibit a higher level of skin conductance responses to familiar versus unfamiliar faces (see e.g. Bate & Cook, 2012).

²⁵ Ellis and Young’s (1990) “mirror image” proposal was originally based on Bruce and Young’s (1986) model of normal face processing. The details of Bruce and Young’s (1986) model have been contested (Breen et al., 2000a; Breen et al., 2001; Bruce & Young, 2012; Ellis & Lewis, 2001; Lewis & Ellis, 2001; Lewis et al., 2001). But the gist of the “mirror image” proposal is widely accepted.

with the Capgras delusion have no difficulty in overtly recognising familiar faces given that they claim that the imposters look like the persons they impersonate, it is an empirically testable prediction that people with the Capgras delusion have abnormal autonomic responses to familiar faces. Since it is proposed, this prediction has been empirically confirmed by five studies (Brighetti et al., 2007; Ellis et al., 1997; Ellis et al., 2000; Hirstein & Ramachandran, 1997; Nuara et al., 2020).

Here it may be worth emphasising that skin conductance responses are not the only form of covert recognition that may be retained in prosopagnosia. Other forms of covert recognition are exhibited by subjects' behavioural performance such as reaction time (Young et al., 1988; Diamond et al., 1994; Barton et al., 2004), and brain activities measured by electroencephalography (EEG) (Bobes et al., 2004; Renault et al., 1989), and by functional magnetic resonance imaging (fMRI) (Liu et al., 2014; Rossion et al., 2003; Simon et al., 2011). If we take it that Ellis and Young's (1990) "mirror image" proposal is not merely talking about skin conductance responses but broadly suggesting that, whereas in prosopagnosia subjects' overt recognition is impaired but their covert recognition is intact, in the Capgras delusion subjects' overt recognition is intact but their covert recognition is impaired, then it may also be predicted that the other forms of covert recognition are impaired in the Capgras delusion. To my best knowledge, it remains to be tested whether other forms of covert recognition that are intact in prosopagnosia are impaired in the Capgras delusion. It is also a remaining question as to how each form of covert or overt recognition may interact with each other to determine one's judgement about people's identities, a complete answer to which should help us better understand, say, to what extent the activity in the autonomic nervous may affect one's judgement about people's identities.

For now, it seems clear that people with the Capgras delusion have reduced autonomic responses to familiar faces. And, compared with pure conjectures about what the anomalous experience in the Capgras delusion is, Ellis and Young (1990) and others' work on the reduced autonomic responses to familiar faces gives us a more concrete ground to find out what the anomalous experience in the Capgras delusion is.

The reduced autonomic responses measured by skin conductance themselves, however, do not suffice to tell us what the anomalous experience is. This is because the activity in the autonomic nervous system is both an unconscious phenomenon (Coltheart, 2005) and "a nonspecific physiological response": the autonomic nervous system can be aroused in many different ways, such as by a loud tone (Breen et al., 2000a, p. 57). In the Capgras delusion, it can be a point of contention about what information is encoded in the reduced autonomic responses, and what kind of anomalous experience is underpinned by the reduced autonomic responses. Regarding the anomalous experience, there are at least three different views in the literature.

First, according to what we can call *the affective view*, which is evident in Ellis and Young (1990), the autonomic activity "carr[ies] some sort of affective tone." (p. 244) That is to say, people with the Capgras delusion experience reduced affective responses at the conscious level. When the person with the Capgras delusion sees his wife, he has "an experience of seeing a face that looks just like ... [his wife], but without experiencing the affective response that would normally be part and parcel of that experience." (Stone & Young, 1997, p. 337)

Our affective states include a wide range of emotions and moods. Among them, the feeling of familiarity is assumed by many scholars to be at the centre of the aetiology of the Capgras delusion. It is also mentioned by Ellis and Young (1990), in which they

sometimes suggest that the autonomic activity carries the tone of the “affective familiarity” in particular. That is to say, people with the Capgras delusion experience a feeling of reduced familiarity when they see familiar faces (see also Young, 2007, 2009). This view is supported by a recent lesion network mapping work by Darby et al. (2017), who analyse the brain lesion locations of a sample of subjects with misidentification delusions including the Capgras delusion, and argue that the lesion locations are functionally connected to a brain region (the left retro-splenial cortex) which is most activated by personal familiar (versus unfamiliar) stimuli in normal subjects. It is, however, worth noting that this finding is contestable: one of the reasons is that the brain lesion locations of some subjects with misidentification delusions do not fit Darby et al.’s analysis (Ferguson et al., 2017; Darby & Fox, 2017). More work is needed to establish that in the Capgras delusion the anomalous experience is a feeling of reduced familiarity.

Second, according to what we can call *the alert view*, advocated by Coltheart (2005), we have an “automatic and unconscious” prediction and comparison system whose job is to make predictions and compare the predictions with inputs, and if the predictions and inputs do not match, an alert will be reported “to consciousness to instigate some intelligent conscious problem-solving behaviour”. In the Capgras delusion, when the person sees his wife, his prediction and comparison system predicts high-level autonomic responses, but this does not match with the received low-level autonomic responses, the mismatch results in the prediction and comparison system’s “reporting to consciousness ‘There’s something odd about this woman.’” (p. 155)

While Coltheart’s proposal about the anomalous experience in the Capgras delusion is based on a detailed analysis of Ellis and Young’s (1990) work on the abnormal autonomic activity, the more

general ideas about the prediction and comparison system²⁶ and the feeling of significance echo Maher's view. Maher (1999) writes:

Survival requires the existence of a detector of changes in the normally regular patterns of environmental stimuli, namely those that are typically dealt with automatically. The detector functions as a general non-specific alarm, a "significance generator," which then alerts the individual to scan the environment to find out what has changed. (p. 558)

Third, according to what we can call *the endorsement view*, the content of the Capgras delusion is already encoded in the anomalous experience (Aimola Davies & Davies, 2009; Davies & Coltheart, 2000; Davies et al., 2001; Davies & Egan, 2013; Fine et al., 2005; Gold & Hohwy, 2000; Langdon & Bayne, 2010; Turner & Coltheart, 2010).²⁷ When the person with the Capgras delusion sees his wife, he has the anomalous experience that "This woman looks like my wife but she is not my wife."²⁸

²⁶ The general idea about the prediction and comparison system is also adopted and further developed by the predictive processing theory of delusions, which emphasises that the prediction error generated by the mismatch between prediction and input is a key factor in the aetiology of delusions (Corlett, 2018; Corlett et al., 2010b; Corlett et al., 2016). The predictive processing theory, however, has been focusing on the subpersonal-level information processing mechanisms, rather than the personal-level phenomena. Given that our present concern is the anomalous experience at the personal level, I shall set aside the predictive processing theory for now. For discussion of the predictive processing theory, see Section 4 of Chapter 4.

²⁷ These papers contain substantial disagreements on other issues: Gold and Hohwy (2000) argue for a one-factor theory (see also Hohwy & Rosenberg, 2005); Davies and Coltheart (2000), Davies et al. (2001), Aimola Davies and Davies (2009), Langdon and Bayne (2010), Turner and Coltheart (2010), and Davies and Egan (2013) argue for versions of the two-factor theory; Fine et al. (2005) is a critique of both.

²⁸ Strictly speaking, there is a fifth view, according to which the transition from the reduced autonomic activity to the delusional belief is an unconscious process and "the first delusion-relevant event of which the patient is aware is the [delusional] *belief* 'That isn't my wife'." (Coltheart et al., 2010, p. 264, emphasis added). For discussion, see McKay (2012), and Davies and Egan (2013).

Let us compare these three views with Maher's conception of anomalous experience. The endorsement view is obviously at odds with Maher's view that delusions are normal explanations, because according to the endorsement view the content of the delusion is already encoded in the anomalous experience and the delusion is hence not an *explanation*, in any ordinary sense, of the anomalous experience. (I will discuss the endorsement view in Section 3.3.3.)

The alert view can be taken as being broadly in line with Maher's conception: both can agree that the subject has an experience that "There's something odd about this woman."

As for the affective view, it is not as straightforward as it may appear whether it should be taken as similar to Maher's conception of anomalous experience. On the one hand, Maher himself seems to take it as a candidate for his conception of anomalous experience when he explicitly mentions Ellis and Young's (1990) work. Maher (1999) writes:

The misidentification syndromes, for example, provide instances in which a focal disturbance of the experience of recognition leads to delusions such as the Capgras, Cotard, and Fregoli syndromes (e.g., Ellis & Young, 1990). Indeed, the model of delusion formation provided in this paper posits that the basic origin lies in the anomalous experience, regardless of how that anomaly arose. (p. 566)

On the other hand, there is, however, an important discrepancy between Ellis and Young's conception of anomalous experience and Maher's conception: while according to Maher (1999) the anomalous experience is "a *vague general feeling* [of significance] that prompts us to look for [an explanation]." (p. 554, emphasis in original), according to Ellis and Young (1990) the anomalous experience has a relatively precise content: that is, the feeling of reduced affective responses or the feeling of reduced familiarity in particular. Of course,

the feeling of reduced affective responses may make the subject experience a feeling of significance. But this feeling of significance is not vague; rather it is relatively definite in that it is caused by the feeling of reduced affective responses.²⁹

Admittedly, the feeling of reduced affective responses may also make the subject wonder what the cause of the reduced affective responses is, and it may appear unclear to the subject what the cause is. But this kind of wondering and unclarity is normally not associated with the “*vague general feeling*” that, according to Maher, would lead to delusions.

To be clear, I am not saying that it is impossible for people with delusions to have the kind of anomalous experience proposed by Maher; the point is that the feeling of reduced affective responses does not seem to be the same as the anomalous experience proposed by Maher; it is obviously not caused by the anomalous experience proposed by Maher; nor does it seem to be the cause of the anomalous experience proposed by Maher.

Now we have to make a choice: we may take it that the vague general feeling of significance is the only conception of anomalous experience proposed by Maher, and accordingly take it that in the above quotation Maher simply gives an inaccurate reading of Ellis and Young’s conception of anomalous experience; or we may take it that Maher actually proposes that both kinds of anomalous experience, i.e.

²⁹ Moreover, while Maher argues that the anomalous experience is the only factor in the aetiology of the Capgras delusion, both Ellis and Young later explicitly argue that a second factor is needed (see e.g. Bell et al., 2006; Stone & Young, 1997). One interpretation of the disagreement is that they are talking about the same kind of anomalous experience, and disagree on whether it is the only factor in the Capgras delusion. I think a careful reading of Maher (1999) and Ellis and Young (1990) suggests that they may not be talking about the same kind of anomalous experience. As an alternative interpretation, we may take it that they are actually talking about different kinds of anomalous experiences. Following the latter interpretation, Maher’s one-factor theory and the two-factor theory are not necessarily incompatible because it is possible that in some cases the delusion is caused by Maher’s anomalous experience as the only factor, whereas in other cases the delusions is caused by Ellis and Young’s anomalous experience and a second factor.

the vague general feeling of significance and the feeling of reduced affective responses, can be the only factor that causes delusions. Neither option seems ideal. Nonetheless, for the convenience of our discussion, in what follows, I will take the second option, and discuss whether either of these conceptions of anomalous experience can be the only factor that causes delusions.³⁰

2.3. A critique of Maher's one-factor theory

Maher's view that the only factor in the aetiology of a delusion is an anomalous experience is often called the one-factor theory. It is

³⁰ The phenomenological approach to delusions has been long focusing on the experiences of people with delusions. However, as summarised by Sass and Pienkos (2013), the phenomenologist is more interested in the abnormal experiences of having delusions rather than the anomalous experiences that give rise to delusions. They write: “[t]he phenomenological approach to delusions focuses on delusion as a *phenomenon* ... The phenomenologist is interested, first and foremost, in understanding *what it is like* to have a delusion, or, more accurately, in understanding the variety of ways in which one might experience delusions and the delusional world.” (p. 632, emphasis in original) Besides, the phenomenological approach has been focusing on delusions in schizophrenia rather than monothematic delusions of neuropsychological origin.

One notable exception is Ratcliffe's (2004, 2008b, 2010) theory of monothematic delusions. Ratcliffe (2010) argues that in the Capgras delusion the anomalous experience has “‘existential changes’ in the structure of experience.” (p. 581) This is an intriguing proposal. But the conception of existential changes does not accord well with the existing empirical evidence. First, the reduced autonomic activity in the Capgras delusion is a focal deficit. It is far from clear how a focal deficit can make existential changes. Second, there are subjects with ventromedial frontal lesions who also have reduced autonomic responses to familiar faces (Tranel et al., 1995). But there is no evidence that these subjects, who have no delusions, have existential changes in the structure of their experiences.

Another important and prevalent idea in the phenomenological tradition is that people with schizophrenia suffer from “changes in perception and automatic processes that are related to the altered experience of self. The altered self-experience, in turn, may be responsible for the emergence of delusions.” (Uhlhaas & Mishara, 2007, p. 142) I acknowledge that there are cases in which an altered self-experience can be a distant causal factor in the aetiology of delusions. However, it is doubtful that this idea is applicable to delusions in general. First, many people with delusions, in particular monothematic delusions of neuropsychological origin, do not seem to suffer from an altered self-experience. Second, even if a person with a delusion does have an altered self-experience, there is still a considerable explanatory gap between the altered self-experience and the delusion. For example, how could an altered self-experience explain the Capgras delusion that his wife is an imposter?

criticised by two-factor theorists, according to whom a second factor is also needed (Coltheart et al., 2011; Davies et al., 2001; Davies & Egan, 2013; Stone & Young, 1997). Maher's view is, however, not without defences (Franceschi, 2010; Noordhof & Sullivan-Bissett, 2021; Reimer, 2009; Sakakibara, 2018; Sullivan-Bissett, 2020).^{31, 32} In this section, I shall discuss the arguments from both parties, and offer a series of arguments suggesting that it is unlikely that the anomalous experience, as it is currently understood in the literature, is the only factor in the aetiology of a delusion.

2.3.1. The intelligibility argument

Jaspers (1913/1997) famously argued that we are unable to “sink ourselves into the psychic situation and *understand ... by empathy*” how a delusion emerges from another mental event (p. 301). One important aspect of Jaspers' point is that there is an absence of an intelligible link between delusions and the mental events that are proposed to be the cause. The absence of an intelligible link can be manifested in many respects (for discussion, see e.g. Davies & Coltheart, 2000; Eilan, 2000; Hoerl, 2019; Walker, 1991). The flouting

³¹ Gunn and Larkin (2019) offer a detailed description of four subjects with delusions and argue that their delusions are “an inevitable consequence of a radical alteration in lived experience.” (p. 151) At first sight, they might appear to offer support to Maher's one-factor theory. However, Maher's one-factor theory is not cited by Gunn and Larkin (2019). I guess the main reason is that Gunn and Larkin (2019) do not discuss whether these subjects have reasoning abnormalities and whether reasoning abnormalities may be the inevitable consequences of a radical alteration in lived experience. If these subjects do have some reasoning abnormality that plays a causal role in the development of their delusions, then Gun and Larkin's account is broadly compatible with the two-factor theory, despite that in these four cases the first factor (a radical alteration in lived experience) may also be the cause of the second factor (see Nie, 2019, for an argument that the two-factor theory does not require a double dissociation between Factor 1 and Factor 2).

³² Gerrans (2002) offers a defence of Maher's one-factor theory of Cotard delusion by arguing that the extreme depression of the subject with the Cotard delusion is sufficient to explain the delusion. However, Gerrans (2014) appears to argue that the extreme depression may also induce a second factor which is “an attributional style that prevents disconfirmation” of the Cotard delusion (p. 126).

of evidence and the specificity of delusions may be among the most striking ones. Can Maher's one-factor theory make delusions intelligible? In particular, can anomalous experience alone fully address the evidence challenge and the specificity challenge?

2.3.1.1. Anomalous experience and the specificity of delusions

As we have seen in the last section, regarding the Capgras delusion, the proposed anomalous experiences include the experience of reduced affective responses or the experience of reduced familiarity (The Affective View), and the experience that "There's something odd about this woman." (The Alert View). Both appear to be able to address the specificity challenge. The reason why the patient's Capgras delusion is only about his wife may be that the reduced autonomic responses and the associated anomalous experiences only occur when he sees his wife's face.

Moreover, it turns out that the reduced autonomic activity and the associated anomalous experiences can explain not only the specific theme that the Capgras delusion is about, but also the specific perceptual mode that many cases of the Capgras delusion appear to be restricted to. For example, Lewis et al. (2001) reported a case in which the patient HL suffered from the voice-specific Capgras delusion.³³ They measured HL's skin conductance responses and found that HL had "normal autonomic responses for faces but reduced autonomic responses for famous voices." (p. 217)³⁴

³³ There are a number of reported cases of the Capgras delusion in which the patients either were blind or suffered from sensory loss (see e.g. Dalgarrondo et al., 2002; Hermanowicz, 2002; Raese & Ibrahim, 2015; Reid et al., 1993; Rojo et al., 1991; Signer et al., 1990). Unfortunately, no measurement of the patients' skin conductance responses was reported in these cases.

³⁴ Hirstein and Ramachandran (1997) reported a case in which the patient DS would believe that "his parents were imposters when he was looking at them but not when speaking to them on the telephone." (p. 437) They measured DS's skin conductance

Recently, Nuara et al. (2020) reported a case in which the Capgras delusion was only about the person's son but not about his daughter, and emerged only when he saw his son but not when he heard his son's voice:

In the summer of 2018, I.F. developed the delusional belief that his son was substituted with an imposter. Such belief was selective for visual modality, i.e., I.F. regarded his son as an imposter only when looking at him, but not while listening to his voice, e.g., while speaking to him over the telephone or from a separate room, without reciprocal visual contact. ... Intriguingly, such delusional belief was not directed towards patient's daughter or other close relatives. (p. 298)

Nuara et al. (2020) measured I.F.'s skin conductance responses and found that in line with the specificity of I.F.'s Capgras delusion, he had reduced skin conductance responses when he saw his son's face but not when he saw his daughter's face; and he had high-level skin conductance responses when he heard either his son's or daughter's voices.

In short, the evidence in the literature clearly shows that the reduced autonomic activity can explain the specificity of the Capgras delusion. This gives us good reason to think that the anomalous experiences related to the reduced autonomic activity can explain the specificity of the Capgras delusion as well. For example, one may argue that I.F. only had a Capgras delusion when he saw his son's face because he only had a feeling of reduced familiarity when he saw his son's face.

responses to familiar versus unfamiliar faces and found that DS has reduced skin conductance responses to familiar faces. But they did not measure DS's skin conductance responses to familiar versus unfamiliar voices.

2.3.1.2. Anomalous experience and the flouting of evidence

What about the evidence challenge? In the Capgras delusion, the patient faces significant counterevidence: the delusion is implausible in the light of common knowledge; it is at odds with the fact that the “imposter” looks like his wife and knows things that only his wife knows; it is also at odds with the testimonies from his friends and relatives; the patient may even have partial insights into the force of the counterevidence and the implausibility of the delusion. Can an anomalous experience underpinned by the reduced autonomic activity explain why the patient obstinately believes that his wife is an imposter?

The imposter hypothesis might be *an* explanation of the anomalous experience in the sense that it can help explain why the patient does not have the normal experience when he sees his wife. But being an explanation itself is not a good reason for the patient to adopt the imposter hypothesis in the light of the significant counterevidence. Moreover, there are lots of other hypotheses which not only are better explanations of the anomalous experience than the imposter hypothesis, but can better deal with the counterevidence. Coltheart (2007), for example, suggests the following two alternative hypotheses:

- The brain-damage hypothesis: “I have suffered a stroke that has disconnected my face recognition system from my autonomic nervous system.”
- The fading-love hypothesis: “I no longer love my wife.”
(p. 1059)

Of course, it is a fact that in the Capgras delusion the patient adopts the imposter hypothesis. But the anomalous experience falls short of explaining why the imposter hypothesis is adopted in the light

of the significant counterevidence and other more plausible hypotheses.

This shortcoming of the anomalous experience suggests either that we may reject the anomalous experience as a candidate factor, and search for some other factor(s) to replace it (Option 1), or that we may accept that the anomalous experience is one candidate factor, but some additional factor or factors must be at play in the aetiology of delusions (Option 2).

I think the explanatory power of the anomalous experience in explaining the specificity of delusions demonstrates that the anomalous experience is indeed an important factor in the aetiology of delusions, and we should, therefore, favour Option 2 over Option 1. That is, some additional factor or factors must be involved in the aetiology of delusions.

2.3.1.3. Two distinct lines of defences of the one-factor theory and their problems

Defenders of Maher's one-factor theory have offered two distinct lines of responses. One is to argue that the imposter hypothesis is not an implausible explanation; the other is to argue that the anomalous experience is much more anomalous than the proposed anomalous experiences: i.e. the experience of reduced affective responses or the experience of reduced familiarity (The Affective View), and the experience that "There's something odd about this woman." (The Alert View). Let us consider them in turn.

Regarding the (im)plausibility of the imposter hypothesis, Reimer (2009) writes:

Although scientists and philosophers (and educated persons more generally) typically regard all such phenomena as equally incredible, that does not mean that they are so regarded by persons generally. Many

people believe in the possibility—if not the actuality—of such spectacular things. A few Google searches of (inter alia) aliens, U.F.O.s, Bigfoot, the Loch Ness Monster, and Elvis confirm as much. It is doubtful that only psychiatric patients are visiting these websites. Thus, the patient's general metaphysical beliefs, the rejection of which is entailed by the Impostor Hypothesis, might not be that deeply entrenched after all. (p. 676)

There are two main points in Reimer's argument. One is that educated persons' general metaphysical beliefs are different from less-educated persons' general metaphysical beliefs: the former is incompatible with the impostor hypothesis while the latter is compatible with the impostor hypothesis. The other is that patients with the Capgras delusion are less-educated persons. If both points are true, then the impostor hypothesis is compatible with the patients' general metaphysical beliefs, and hence is not implausible for the patients.

I guess it might strike someone as an intuitive view that if a person is less educated or has some abnormal metaphysical beliefs, then they are somewhat more likely to be diagnosed with delusions. But this does not entail that people with delusions, in general, are less educated or have abnormal metaphysical beliefs. There is no statistical evidence in support of the view that people with the Capgras delusion are less educated (for reviews, see Pandis et al., 2019; Currell et al., 2019; Bell et al., 2017; Edelstyn & Oyebode, 1999). Nor is there any statistical evidence that the general metaphysical beliefs or knowledge possessed by people with delusions are significantly different from that possessed by people without delusions. In fact, when common knowledge is explicitly examined, there is evidence that people with delusions do not have abnormal knowledge. For example, Young and Leafhead (1996) tested a patient WI who had the Cotard delusion that he was dead. They found that

WI's knowledge about being alive or dead did not differ from that of non-delusional control subjects:

To explore WI's knowledge of whether other people were alive or dead, he was given a simplified version of Kapur's Dead or Alive Test (Kapur et al., 1989). The names of 30 famous people were presented, 10 of whom were alive at the time of testing, with the remaining 20 having died between 1960 and 1989. In each case, WI was asked whether the person was alive or dead. For those he thought had died, he was also asked to give the year in which they died (scored as correct if it belonged to the appropriate decade) and whether they were killed or died of natural causes. [It turned out that] ...the performance of WI and eight control subjects matched for age and education (mean age 28 years, SD 2.92 years). For all parts of the test, WI scored as well as the controls. Hence, it is clear that his delusional belief in his own death had not affected WI's general knowledge of whether other people were alive or dead, and was not accompanied by any retrograde amnesia. (p. 156)

Moreover, the fact that some patients have partial insights into the force of counterevidence and the implausibility of their delusions suggests that their general knowledge is not significantly different from non-delusional people (Section 2 of Chapter 1).

Let us turn to the arguments appealing to the anomaly of the anomalous experience. Maher (1999) argues that "the kinds of anomalous experience that deluded patients have appeared to be much more *intense* and *prolonged* than those that occur to the population in general." (p. 566, emphasis added). Defenders of Maher's one-factor theory have repeatedly emphasised this point. Reimer (2009), for example, writes: "[T]he experience is widely thought to involve a *profound* feeling of unfamiliarity, of estrangement.

This is a feeling that goes well beyond a mere absence of familiarity.” (p. 674) However, Maher and Reimer did not tell us about what exactly the intense, prolonged, profound anomalous experience is. Merely insisting that there is such an anomalous experience that is the only factor in the aetiology of delusions offers limited help in explaining delusions (see also Davies et al., 2001, pp. 146-147).

Sakakibara (2018) offers an illustration of the intense anomalous experience by comparing it to intense emotions:

[The intense anomalous experiences] are functionally comparable to intense emotions that we sometimes have, in that both work as “irruptive motivation.” Strong emotions irrupt into deliberative means-end reasoning and take over one’s judgements and actions (Prinz, 2004). For instance, intense fear may cause one to “run away from situations that could be rewarding” (Prinz, 2004, p. 84). Similarly, an intense experience irrupts through deliberative reasoning and rushes the subject into holding delusional thoughts, even though this is not warranted given the totality of evidence. (p. 177)

Suppose it is true that an intense anomalous experience would rush the subject’s belief forming processes. But if the subject is in a rush, wouldn’t it be more likely that they would rush into some more plausible and available hypotheses, as opposed to the far-fetched delusional hypotheses? In the Capgras delusion, wouldn’t it be more likely that the patient rushes to believe the brain-damage hypothesis or the fading-love hypothesis, as opposed to the imposter hypothesis?

Moreover, Sakakibara (2018) seems to indicate that when a person is rushed into holding a delusion, the counterevidence is excluded from the person’s consideration of what to believe. But this does not explain why delusions persist even when the counterevidence is clearly presented to the person and the person

appears to have partial insights into the force of the counterevidence and the implausibility of their delusions. It may be true that intense fear may cause one to run away from situations that could be rewarding if they stay, but once it is shown to the subject that the situations are indeed rewarding when they stay, it is unlikely that they would keep running away. If the intense anomalous experience were like the intense fear, then once the counterevidence is clearly shown to the patients, it is unlikely that they would keep the delusions. The fact that people with delusions fail to reject their delusions in the light of counterevidence, in turn, suggests that Sakakibara's conception of the intense anomalous experience falls short of addressing the evidence challenge of delusions.

To my best knowledge, there has been no convincing account of the anomalous experience such that the anomalous experience itself suffices to explain why delusions are formed and maintained in the face of counterevidence. Of course, the lack of such an account, so far, does not prove that it is impossible that there exists such a kind of anomalous experience in the light of which delusions are inevitable. But for the one-factor theory to transform from being a conjecture that has not proven impossible to being a highly plausible account, we need to know more about what the anomalous experience is.

In my early work, I sketched an alternative way to help us have a grasp on what the anomalous experience could be (Nie, 2017). I argued that it is *theoretically* possible that there might exist a certain kind of anomalous experience, of which the imposter hypothesis is the best explanation provided by our cognitive system with limited processing power.

To see the point, let us first consider a thought experiment by Locke (1690/1975):

A studious blind man, who had mightily beat his head
about visible objects, and made use of the explication

of his books and friends, to understand those names of light and colours, which often came in his way, bragged one day, that he now understood what scarlet signified. Upon which his friend demanding, what scarlet was? the blind man answered, it was like the sound of a trumpet. (Chapter IV, §11)

The intuition behind Locke's thought experiment is that a certain kind of experience is necessary for a certain kind of knowledge (Campbell, 2002). It seems to me that the reverse of this intuition may also be right. That is, a certain kind of knowledge and the corresponding cognitive abilities may be necessary for properly explaining, understanding, or even describing a certain kind of experience. If this is correct, then we may have the following thought experiment based on Locke's:

- I. Imagine Jane who, like Locke's blind man, is colour blind.
- II. But, unlike Locke's blind man, Jane lives in a world in which no one has the sensation of colour and there are no colour concepts and words.
- III. One day, Jane somehow has a faint sensation of colour. She looks at the scarlet wall, and notices that something is different. But, without the help of any colour concepts and words, Jane is not able to know that it is the scarlet colour that causes the difference.³⁵

³⁵ We can distinguish between primary experiences and cognitive experiences (for a similar distinction, see also Dewey, 1929; Eames, 1964). Primary experiences are non-cognitive and prior to a conceptual and linguistic transformation. When we see an apple, there is a primary experience of the apple even if we have no concept and word about apple. Similarly, Jane in the thought experiment has a primary experience of colour even though she has no concepts and words about colour. As for the cognitive experience, since we have the concept and word about apple, we also have a cognitive experience of the apple. By contrast, Jane does not have a cognitive experience of colour.

- IV. Suppose the experience of the scarlet wall causes other mental or physical events in Jane, which are, in a sense, similar to the events caused by the experience of her listening to a trumpet.
- V. For Jane, this anomalous experience may give rise to a belief that the wall secretly plays trumpet music to her.
- VI. In her world, Jane may be diagnosed as having a play-trumpet-wall delusion.

If this thought experiment is convincing, then in another similar thought experiment Jane could be just like us but travels to a different world, say, Mars, which gives her an anomalous experience that lacks palpable analogies in her previous experiences. Or Jane may not need to live in a colour-blind society or travel to Mars to have anomalous experiences. Her anomalous experience might simply be due to some abnormalities in her perceptual systems. If this analysis is along the right lines, then what happens to Jane in this last scenario might be quite close to what happens to people with delusions in the real world.

It is, however, important to note that even if it is *theoretically* possible that there might exist a certain kind of anomalous experience such that the imposter hypothesis is the best explanation our cognitive system can offer, this account still does not tell us what the anomalous experience is. It is also unclear whether this is what happens in the Capgras delusion. If there is some empirical evidence that the information received by the patients overloads their cognitive systems, then we may be more confident in the account I just outlined. But I am not aware of such evidence.

To conclude, despite being able to help explain the specificity of delusions, the anomalous experience alone falls short of explaining why the imposter hypothesis is adopted in the light of the significant counterevidence and other more plausible hypotheses. Some additional factor or factors are, therefore, needed.

2.3.2. The dissociation argument: three variants

The standard argument offered by two-factor theorists against Maher's one-factor theory is the dissociation argument (see e.g. Coltheart & Davies, 2021a, Table 1; Coltheart et al., 2011). We can discern three variants of the dissociation argument.

The first variant argues that there are cases in which the individuals have an anomalous experience similar to the proposed anomalous experience that gives rise to a delusion, but they do not have the delusion. A second factor is, therefore, needed to explain the delusion. Turner and Coltheart (2010) reported a case (studied by Nora Breen and Mike Salzberg) in which the person had an anomalous experience similar to the anomalous experience of people with the Capgras delusion, but the person did not have the Capgras delusion:

Interviewer: After the operation, did you notice any change?

Patient: Yes, the first thing I noticed was Mum, when she walked in the room, it was just a . . . it was cardboard . . . an image, if that's the right word . . . it was like a picture of her, but it wasn't her . . . it's hard to, I can't sort of explain it 'cause it . . . inside . . . it's hard.

Interviewer: Right. It looked like a real person, did it?

Patient: Oh, yeah. There's Mum walking in the room.

Interviewer: Right. So you knew it looked like your Mum, but . . .

Patient: Whether personality or what I don't know, but she was different, something was different about her

Interviewer: Did she look different or . . . ?

Patient: She . . . I don't know . . . so you can look different by expression or . . . just through feeling, you can look different by, you know, doing your hair or

whatever, but it wasn't different in that way, it was of having the different inside of her, I can't sort of explain it

Interviewer: Right is it that you felt different towards her?

Patient: I wouldn't have a clue because I'm just laying there and Mum's walked in the room . . . well, this picture of Mum . . . and started talking but it was only a picture of her but it didn't feel like her.

Interviewer: Right. OK.

Patient: Just didn't feel like her.

Interviewer: How did that feel to you?

Patient: Scary Has their lifestyle changed? Have I changed? Have they changed in a funny sort of way? I don't know. It's weird and it gets confusing. (pp. 371-372)

The second variant argues that there are cases in which the individuals have a neuropsychological deficit similar to the neuropsychological deficit underpinning the anomalous experience that gives rise to a delusion, but they do not have the delusion. A second factor is, therefore, needed to explain the delusion (see e.g. Davies & Egan, 2013, p. 691). Regarding the Capgras delusion, there were individuals with ventromedial frontal lesions who had reduced autonomic responses to familiar faces and probably had an anomalous experience similar to that in the Capgras delusion, but they did not have the Capgras delusion (Tranel et al., 1995).³⁶

³⁶ As discussed in Section 2.2, there are different interpretations of the reduced autonomic responses to familiar faces, and in particular the relationship between the reduced autonomic responses and the anomalous experience. Even if both the individuals with ventromedial frontal lesions and people with the Capgras delusion have reduced autonomic responses to familiar faces, this does not entail that they have a similar anomalous experience. Nor does the same reduced autonomic responses entail that the cause of the reduced autonomic responses in the individuals with ventromedial frontal lesions is the same as the cause of the reduced automatic responses in people with Capgras delusions.

Ellis and Lewis (2001) proposed an alternative interpretation, according to which the individuals with ventromedial frontal lesions have an intact affective experience of

Defenders of Maher’s one-factor theory may argue that the anomalous experience of individuals without delusions significantly differs from the “anomalous” experience of people with delusions (Franceschi, 2010; Reimer, 2009; Sakakibara, 2018; Sullivan-Bissett, 2020). Again, I acknowledge that it is a theoretical possibility. But such a possibility has not been supported by empirical evidence. The more realistic suggestion is that a second factor is needed.

The third variant of the dissociation argument does not directly argue against Maher’s view that an anomalous experience is the only factor in the aetiology of a delusion. Instead, it argues that the proposed neuropsychological deficit, which may or may not lead to an anomalous experience, is not sufficient. Regarding the Capgras delusion, it argues that the reduced autonomic activity itself is not sufficient because there are individuals who have the reduced autonomic activity but do not have the Capgras delusion (Tranel et al., 1995). Compared to the second variant, the third variant does not make a claim about whether there is a shared anomalous experience between individuals with and without delusions. An exemplar of the two-factor theory that makes no claim about the existence of a conscious anomalous experience is proposed by McKay (2012). His theory allows the possibility that both the first factor and the second

familiar faces, whereas people with the Capgras delusion have a reduced affective experience of familiar faces. However, both proposals face their own problems. Regarding people with the Capgras delusion, we have seen that it is not clear that the reduced autonomic activity carries an affective tone (Breen et al., 2000a; Coltheart, 2005). Regarding the individuals with ventromedial frontal lesions, there is some evidence suggesting that they have impaired affective experiences (see Fine et al., 2005).

While more studies are needed to help us understand the nature of the reduced autonomic activity and the anomalous experience, it is worth noting that the dissociation argument does not rely on a single dissociation case but is supported by a wide range of dissociation cases regarding various delusions (McKay, 2019). For a collection of dissociation cases regarding Capgras delusion, Fregoli delusion, Cotard delusion, mirrored-self misidentification, somatoparaphrenia, and alien control delusion, see Coltheart et al. (2011) and Coltheart and Davies (2021a, Table 1).

factor are unavailable to consciousness (McKay, 2012, footnotes 3 & 7).

All three variants of the dissociation argument suggest that the proposed first factor, either the neuropsychological deficit or the anomalous experience, is not sufficient, and a second factor is needed. But, importantly, the dissociation arguments do not entail that the second factor in need is a reasoning abnormality (Nie, 2016, 2017). Nor do they preclude the possibility that there are more than one missing factors. As far as the dissociation arguments are concerned, it is also possible that the missing factor in need is not a reasoning abnormality but some other kind of neuropsychological abnormality. In Chapters 4 and 5, I shall propose a factor that is not a reasoning abnormality. For now, I shall focus on the proposal by two-factor theorists that the missing factor is a reasoning abnormality (Aimola Davies & Davies, 2009; Aimola Davies et al., 2009; Breen et al., 2000a; Coltheart, 2007, 2010; Coltheart et al., 2011; Davies & Coltheart, 2000; Davies et al., 2001; Davies & Egan, 2013; Langdon & Coltheart, 2000).

A not very uncommon misunderstanding of the dissociation argument is that it requires a *double* dissociation between perceptual anomaly and reasoning abnormality (e.g. Connors & Halligan, 2020; Corlett, 2019; Corlett & Fletcher, 2021). As I have explained, the dissociation argument only shows that the proposed first factor is not sufficient because of the *single* dissociation cases in which the proposed first factor is present, but the corresponding delusion is not. The second factor is proposed to help explain the single dissociation cases. The dissociation argument does not, and does not need, to make a further claim about the relationship between the nature of the first factor and that of the second factor: for example, it does not make a claim about whether there is some neuroanatomical overlap between them, whether there is some overlap between the cognitive underpinnings of the first factor and those of the second factor, or

whether there is some degree of interaction between the first and second factors, let alone a double-dissociation relationship between them (see also Davies et al., 2001, footnote 15; for a discussion of other misunderstandings of the two-factor theory, see Nie, 2019).

2.3.3. The empirical evidence of reasoning abnormalities

Maher's one-factor theory was supported by the repeated failure of the empirical studies to find logical reasoning abnormalities in schizophrenia (Mujica-Parodi et al., 2000; Maher, 2001). Since then, however, more studies have been developed and a variety of reasoning abnormalities have been found to be associated with delusions. In a recent review of empirical studies, McLean et al. (2017) argue that delusions are associated with the jumping to conclusions bias, the bias against disconfirmatory evidence, the bias against confirmatory evidence, and liberal acceptance. Based on their own empirical studies of anosognosia (with which the subject may have the delusion that their paralysed arm is not paralysed but normal) and the review of the literature, Aimola Davies and colleagues argue that people with delusions suffer from an impairment of working memory and/or executive function, which are essential for reasoning processes (Aimola Davies et al., 2009; Aimola Davies & Davies, 2009). These new empirical findings suggest that some reasoning abnormalities may be an important factor in the aetiology of delusions.

Defenders of the one-factor theory do not deny the validity of these empirical studies. Instead, they argue that the reasoning abnormalities in delusions are nonetheless within the normal range. To see their points, we need to take a closer look at their arguments. In the following, I will discuss what I will call the meta-theory argument and the analogical argument in defence of the one-factor theory.

2.3.3.1. The meta-theory argument in defence of the one-factor theory and its problems

In her defence of Maher's one-factor, Sullivan-Bissett (2020) argues that people with delusions "have normal-range reasoning applied to abnormal experiences." (p. 683) What is a "normal-range" reasoning process? How should we distinguish a normal-range reasoning process from an abnormal-range reasoning process? Sullivan-Bissett does not offer an answer. Instead, she argues that the burden is equally on two-factor theorists "since [they argue] that there is some cognitive feature of subjects who have delusions which is abnormal and differentiates them from the non-delusional population." (p. 684)

From the perspective of two-factor theorists, there is a clear difference between how people with delusions reason and how people without delusions reason, and the reasoning in delusions falls outside of the normal range. However, Sullivan-Bissett is correct that without a meta-theory of what counts as a normal-range reasoning process, one-factor theorists could insist that the reasoning in delusions still "fall[s] into the normal range" (p. 683). Call this the meta-theory argument.

How to break the impasse? One way, of course, is to develop a meta-theory of what counts as a normal-range reasoning process. I shall not take this approach. Instead, I argue that appealing to the meta-theory argument would do more harm than benefit to the one-factor theory.

To my knowledge, there is no meta-theory of what counts as an anomalous experience. So, following a similar meta-theory argument, a zero-factor theorist can argue that the anomalous experience falls within the normal range and hence we should reject the one-factor theory as well.

There is no widely accepted meta-theory of what counts as a delusional belief either.³⁷ So, following another similar meta-theory argument, a delusion-denialist (compare Szasz, 1974) could argue that delusional beliefs fall within the normal range, and hence there is no need to develop any theory of delusions in particular: the existing theories of normal beliefs are sufficient for explaining delusions.

Here is a dilemma for one-factor theorists. On the one hand, if they retain the meta-theory argument, then they can defend the one-factor theory against the two-factor theory, but they have to explain why the zero-factor theorist and the delusion-denialist are wrong. It is far from clear how this can be done because the zero-factor theorist and the delusion-denialist are using similar meta-theory arguments. On the other hand, if they give up the meta-theory argument, then they do not have to deal with the zero-factor theorist and the delusion-denialist, but their defence of the one-factor theory will be lost as well.

For now, I shall set aside the meta-theory argument, and take it that the empirical evidence clearly shows that delusions are associated with reasoning abnormalities, some of which may be a second factor in the aetiology of delusions.

2.3.3.2. The analogical argument in defence of the one-factor theory and its problems

By appealing to Cassam and others' work on conspiracy theories, Noordhof and Sullivan-Bissett (2021) offer an analogical argument that the irrationality in delusions is within the normal range. They write:

³⁷ Interestingly, this appears to be the view held by Sullivan-Bissett too. In a paper co-authored by Bortolotti, Sullivan-Bissett, and Gunn (2016), they write: "[T]he prospect of arriving at a principled way to distinguish delusional from non-delusional beliefs is not promising." (p. 48)

Consider also conspiracy theories, which we will understand as explanations of events that appeal to the intentional states of conspirators, who intended the event and kept their intentions and actions secret (Mandik, 2007, p. 206). Those who believe in such theories—so-called conspiracy theorists—are prime examples of epistemically irresponsible subjects whose beliefs seem utterly impervious to counterevidence. As Quassim Cassam points out, ‘there aren’t too many examples of committed conspiracy theorists changing their minds’ (2019, p. 93). Conspiracy theorists are especially relevant to discussion here since, perhaps similarly to some monothematic delusions, ‘[t]here is almost no explanation that isn’t too bizarre for the conspiracy theorist’s taste’ (Cassam, 2019, p. 22). ... many researchers interested in conspiracy theorists make no claims about clinically abnormal cognition, rather, they appeal to individual differences in personality to explain being conspiracy-minded (see Cassam, 2019, pp. 40–43 for discussion) or as involving a particular worldview (Keeley, 1999, p. 123, Cassam, 2019, p. 100). Such normal range irrationality is the kind of thing which can contribute to such thinkers displaying epistemic irresponsibility. ... Similar things can be said for folk with monothematic delusions. ... the epistemic irresponsibility displayed by such subjects is representative of normal range irrationality...

Their argument seems to go as follows: The kind of epistemic irresponsibility in delusions is the same as the kind of epistemic irresponsibility in conspiracy theories; normal-range irrationality, by which Noordhof and Sullivan-Bissett (2021) mean, e.g. the differences in personality, can explain the kind of epistemic irresponsibility in

conspiracy theories; therefore, normal-range irrationality can explain the kind of epistemic irresponsibility in delusions.

While it is true that both delusions and conspiracy theories are notorious types of beliefs that are impervious to counterevidence, and in this very general sense the subjects in both cases are epistemically irresponsible, it does not follow that the kind of epistemic irresponsibility in delusions is the same as the kind of epistemic irresponsibility in conspiracy theories. There is a wide range of types of beliefs that are impervious to counterevidence, which, besides delusions and conspiracy theories, may also include akratic beliefs (Borgoni & Luthra, 2017; Heil, 1984; Hookway, 2001; Mele, 1986), superstitious beliefs (Scheibe & Sarbin, 1965), and some of the religious beliefs (Jones & Watson, 2018; McKay & Ross, 2020). In spite of the fact that these beliefs are all impervious to counterevidence, and in this very general sense the subjects with any of these beliefs may be taken as being epistemically irresponsible, it is, however, unlikely that the specific kinds of epistemic irresponsibility in all these beliefs are the same.

One might defend Noordhof and Sullivan-Bissett (2021) by arguing that even though different kinds of epistemic irresponsibility are involved in different types of beliefs that are impervious to counterevidence, we can still attribute the same normal-range irrationality to all of them. This defence is, however, too weak. This is because if one accepts that the kind of epistemic irresponsibility in delusions differs from the kind of epistemic irresponsibility in conspiracy theories, then it is unclear what the reason is for thinking that the kind of irrationality in delusions is still within the normal range.

We might have some reason for thinking that the same normal-range irrationality could explain delusions if it turned out that the specific way delusions are impervious to counterevidence is the same as the specific way conspiracy theories are impervious to counterevidence. A closer comparison between Cassam and others'

account of conspiracy theories and what we already know about delusions (Chapter 1), however, shows that the latter is not true.

There are many significant differences between the epistemic irresponsibility in conspiracy theories and the epistemic irresponsibility in delusions. For example, conspiracy theories are often politically, financially, or ideologically motivated, whereas delusions are not; Conspiracy theories are based on the premodern view that “complex events are capable of being controlled by a small number of people acting in secret, and that this is what gives these events a deeper meaning” (Cassam, 2019, p. 26; see also Keeley, 1999), whereas delusions lack such a base. Conspiracy theories are often shared by a group of individuals or within a certain community, whereas there is no “community united in common delusions” (Jaspers, 1913/1997, p. 284). These differences already suggest that the irrational factors in delusions are much more severe than the so-called normal-range irrationality in conspiracy theories.

Furthermore, when evidence is concerned, Cassam (2019, Chapter 4; 2020) argues that conspiracy theories are impervious to counterevidence in a *self-sealing* way (see also Sunstein & Vermeule, 2009). Cassam (2020) writes:

[T]he arguments that give rise to them [i.e. conspiracy theories] are designed to be immune to refutation. Contrary evidence is attributed to the conspiracy and the absence of evidence of conspiracy is taken as evidence of the skill of the conspirators. (p. 3)

For example, for conspiracy theorists who believe that the Bush administration is behind the 9/11 attacks, the official report of the 9/11 Commission is part of the conspiracy, and the lack of evidence that the Bush administration is behind the 9/11 attacks is the evidence that the Bush administration is very good at hiding their conspiracy.

Are delusions impervious to counterevidence in such a self-sealing way? It does not seem so. It is not a characteristic of delusions that the subjects take the absence of evidence as evidence that their delusions are true. Nor is it a characteristic of delusions that the subjects attribute contrary evidence to their delusions. Rather, they can often recognise contrary evidence as contrary evidence, even though they do not change their delusions in the face of contrary evidence (see Chapter 1; Corlett & Fletcher, 2021, p. 233).

To be clear, I do not deny that there may be cases of delusions that display one or some of the self-sealing features of conspiracy theories. For example, regarding the Capgras delusion, some theorists suggest that the evidence that the “imposter” looks exactly like the patient’s wife and knows things only his wife knows might be taken as the evidence that the “imposter” is good at impersonating his wife. It is, however, not clear whether this is the typical way patients with the Capgras delusion treat evidence. In any case, even if these are some cases in which the patients treat evidence in this way. This still does not establish that being self-sealing is a general feature of delusions.

The self-sealing feature of conspiracy theories indicates that the evidence that is at odds with conspiracy theories is, in a sense, explained away in a logically coherent, albeit unsound, way. By contrast, the evidence that is at odds with delusions is not explained away: instead, delusions are held in the face of counterevidence. This significant difference between the way conspiracy theories are impervious to counterevidence and the way delusions are impervious to counterevidence suggests that, even though the kind of irrational factors in conspiracy theories is within the normal range, it does not follow that the irrational factors in delusions are within the normal range. On the contrary, the flagrant ways delusions flout evidence suggest that if some irrational factors, such as the reasoning abnormalities that are empirically demonstrated to be associated with

delusions, play important roles in the aetiology of delusions, then they are likely to be beyond the normal range.

So far in Section 2, I have introduced the anomalous experience hypothesis by focusing on Maher's one-factor theory and argued that the anomalous experience can help us address the specificity challenge of delusions, but have difficulties in addressing the evidence challenge. Moreover, I have also summarised several existing arguments and developed a few new arguments, all of which suggest that there is, at least, a second factor, for which reasoning abnormalities are a candidate, in the aetiology of delusions. In the next section, I shall discuss to what extent reasoning abnormalities can help explain delusions.

3. The reasoning abnormality hypothesis

In the literature, there are three influential groups of theories in which certain reasoning abnormalities play an important role in the aetiology of delusions. The first is the two-factor theory (Aimola Davies & Davies, 2009; Aimola Davies et al., 2009; Breen et al., 2000a; Coltheart, 2007, 2010; Coltheart & Davies, 2021a; Coltheart et al., 2011; Davies & Coltheart, 2000; Davies et al., 2001; Davies & Egan, 2013; Langdon & Coltheart, 2000); the second is what we can call the cognitive-bias theory, according to which delusions are products of cognitive biases, such as the jumping to conclusions bias (Garety & Freeman, 2013; Garety et al., 1991; Ward & Garety, 2019); The third is the predictive processing theory (Corlett, 2018; Corlett et al., 2007; Corlett et al., 2009; Corlett et al., 2010b; Corlett & Fletcher, 2015; Corlett et al., 2016). I shall discuss the predictive processing theory in Chapter 4. The focus of this section will be on the question of to what extent the cognitive-bias theory and the two-factor theory can help us address the evidence and specificity challenges of delusions.

3.1. The cognitive-bias theory: jumping to conclusions or jumping to delusions³⁸

The study of cognitive biases has been one of the most proliferating areas in experimental psychology (Kahneman, 2011; Kahneman et al., 1982). The prevalence of cognitive biases in ordinary people like you and me is now well recognised. Very roughly, to say that ordinary people suffer from a certain kind of cognitive biases in a certain context is to say that the way ordinary people make judgements in that context deviates from the way an ideally rational agent would make their judgements in the same context. In this sense, the cognitive biases explain the divergence of our ordinarily biased beliefs from the ideally rational agent's unbiased beliefs. There is a striking divergence of delusional beliefs from both ordinarily biased beliefs and unbiased beliefs (see Chapter 1). It appears to be an intriguing idea that certain kinds of cognitive biases may help explain the divergence of delusional beliefs.

This idea was empirically supported by a seminal experiment in which, compared to the participants without delusions, the participants with delusions show what has been called the jumping to conclusions (JTC) bias (Huq et al., 1988). Since then, the JTC bias as well as quite a few other cognitive biases have been shown to be associated with delusions (for a review, see McLean et al., 2017). A modest interpretation of the association is that the cognitive mechanisms and/or neuroanatomies underpinning these cognitive biases overlap those underpinning delusions. By contrast, a more ambitious and prevalent interpretation is that these cognitive biases

³⁸ The phrase *jumping to delusions* is first used by McKay et al. (2007) in the narrow sense that delusions are formed in virtue of the JTC bias tested by the beads task. Here I use the term in a broader sense by which I mean that patients somehow jump to delusions, but I do not have the commitment that the JTC bias is the reason why patients jump to delusions. Instead, in this section I will argue that the JTC bias is unlikely to be the main reason why patients jump to delusions.

themselves are causal factors in the aetiology of delusions (Ward & Garety, 2019; Garety & Freeman, 2013; Garety et al., 1991). For example, Ward and Garety (2019) write: “JTC plays a causal role in delusion development and maintenance”. (p. 81)

Since Huq et al.’s (1988) work on the JTC bias, there have been concerns about whether the proposed cognitive biases are associated with delusions in general or associated with only some types of delusions (e.g. Diaz-Cutraro et al., 2021), or associated with some other deficits of the participants (for reviews, see e.g. Dudley et al., 2016; Garety & Freeman, 2013; Livet et al., 2020; McLean et al., 2017; Ross et al., 2015; So et al., 2016). For example, a recent experiment suggests that the JTC bias “is likely to reflect inattentive responding, rather than the presence of floridly delusional participants.” (Sulik et al., 2021) Here I will not engage with this debate. Instead, I will assume that some cognitive biases, in particular the JTC bias, are associated with delusions. My aim is to elaborate on what I think is a fundamental problem with the cognitive-bias theory of delusions in general. It concerns what I take to be a significant drawback of the experimental paradigms in which the proposed cognitive biases are measured. Since the nature of a cognitive bias is determined by the experimental paradigm in which it is measured, I argue that the drawback of the experimental paradigm imposes an important limitation on the explanatory power of the cognitive bias. I shall illustrate the point by focusing on the experimental paradigm in which the JTC bias is typically measured: the beads task (Huq et al., 1988).

In the beads task, participants are informed that there are two jars of beads:

- Jar A contains 85 pink and 15 green beads.
- Jar B contains 15 pink and 85 green beads.

These two jars are then hidden from the participants and one jar is chosen by the experimenter. Beads are drawn from the chosen jar one at a time and are shown to the participants. At each draw, the participants are required to either make a conclusion about which jar the beads are from or ask for another draw until a conclusion is made. The number of draws for the participants to reach a conclusion is noted. The result is that participants with delusions ($M=1.22$, $SD=1.57$) require fewer draws than psychotic participants without delusions ($M=3.58$, $SD=3.51$) and normal participants ($M=2.6$, $SD=1.17$).

This result has been taken as showing that people with delusions have the JTC bias. But, as an empirical experiment, the beads task is open to interpretations along many dimensions. What exactly the JTC bias means depends on how the beads task is interpreted. Under an extremely broad interpretation, the beads task shows that people with delusions are more prone to reasoning errors than people without delusions, and the JTC bias is understood as broad as including all erroneous reasoning processes. In this sense, the JTC bias is a cause of delusions as long as erroneous reasoning processes are a cause of delusions. Obviously, this interpretation is too broad and does not tell us much about what exactly is erroneous in delusions. Under an extremely narrow interpretation, the beads task demonstrates nothing more than that people with delusions make hasty decisions on the basis of fewer *beads* than people without delusions. In this sense, the JTC bias is not a cause of delusions because delusions are not formed as responses to *beads*. Obviously, this interpretation is too narrow.

Of course, no one holds the extremely broad or extremely narrow interpretations. The prevalent interpretation is that the JTC bias is “mak[ing] decisions on the basis of limited evidence” (Ward & Garety, 2019, p. 80). Nonetheless, it seems to me that this interpretation lies somewhere close to the extremely broad interpretation. There are a number of different ways one makes (hasty)

decisions on the basis of limited evidence, and there are many kinds of beliefs that are formed on the basis of limited evidence: such as conspiracy theories (Cassam, 2019; Keeley, 1999; Sunstein & Vermeule, 2009), akratic beliefs (Borgoni & Luthra, 2017; Heil, 1984; Hookway, 2001; Mele, 1986), superstitious beliefs (Scheibe & Sarbin, 1965), and some of the religious beliefs (Jones & Watson, 2018; McKay & Ross, 2020). If the JTC bias is interpreted as referring to all of these different ways as a whole, then the JTC bias might be taken as a cause of delusions *if* delusions are one kind of beliefs that are formed on the basis of little evidence. But, in this broad sense, the JTC bias falls short of explaining why someone jumps to delusions in particular, rather than jumps to conspiracy theories, akratic beliefs, superstitious beliefs, and so on.³⁹

A less broad and more fitting interpretation, I think, is that the beads task demonstrates a unique way one makes decisions in the light of limited evidence. Under this interpretation, the extent to which the JTC bias can help explain delusions depends on the extent to which this unique way is similar to the way delusions are formed.

3.2. A critique of the cognitive-bias theory

What is the unique way people with delusions make decisions in the beads task? In the beads task, although people with delusions require less evidence to reach a conclusion, the evidence (say, one pink bead) still provides greater support to their conclusion (Jar A with more pink beads) than the alternative conclusion (Jar B with more

³⁹ Some critics argue that the “portrayal of people with delusions as having a JTC bias is a redescription rather than an explanation.” (Corlett & Fletcher, 2014, p. 399) This critique seems to go too far, because saying that people with delusions as having a JTC bias is only one of many possible explanations of why delusions are held in the light of little evidence. For many scholars who are more sceptical about the role of the JTC bias in the aetiology of delusions (see e.g. Aimola Davies & Davies, 2009; Fine et al., 2007), the portrayal of people with delusions as having a JTC bias is certainly not a redescription of what is going on in delusions.

green beads). More importantly, after seeing a pink bead, people with delusions do not jump to bizarre conclusions, such as that the bead is from Jar C (it is bizarre because there is no Jar C in the experiment), or from Mars, or from nowhere. If people made such bizarre conclusions, they would not (only) suffer from the JTC bias. That is to say, the JTC bias can help explain why a person jumps to conclusions like that the pink bead is from Jar A, but it falls short of explaining why a person jumps to bizarre conclusions, such as that the bead is from Jar C, or from Mars, or from nowhere.

Can the JTC bias explain why one ends up with delusions? The answer depends on whether the relationship between the evidence possessed by people with delusions and their delusions is more like the relationship between the pink bead and the conclusion of Jar A, or the relationship between the pink bead and the bizarre conclusions, such as that the bead is from Jar C, or from Mars, or from nowhere.

Let us consider the evidence possessed by people with delusions. In the Capgras delusion, the anomalous experience may be the experience of reduced affective responses or the experience of reduced familiarity (The Affective View), or the experience that “There’s something odd about this woman.” (The Alert View) If the anomalous experience is the person’s evidence in response to which they form the Capgras delusion, then the relationship between the evidence and the Capgras delusion seems more like the relationship between seeing a pink bead and the bizarre conclusions such as the bead is from Jar C, or from Mars, or from nowhere.⁴⁰ Since the JTC bias falls short of explaining why one jumps to bizarre conclusions, it falls short of explaining why one jumps to the Capgras delusion too.

What does this tell us about the possible role of the JTC bias in the aetiology of delusions? In an influential critique of the JTC bias,

⁴⁰ If the subject does not form a delusion but merely forms a belief like that “I no longer love my wife” (Coltheart, 2007, p. 1059), then the JTC bias might be a cause of this belief.

Fine et al. (2007) argue that “the JTC bias is not relevant to the formulation of delusional hypotheses. ...[And] the JTC bias seems to have no role to play in the maintenance of delusion in the face of disconfirmatory evidence.” (p. 74; see also Ashinoff et al., 2021) Fine et al.’s key point appears similar to what I have discussed: that is, in the light of the evidence including the subject’s anomalous experience and the significant counterevidence, the JTC bias falls short of explaining why the subject jumps to delusions. However, I do not think that this supports Fine et al.’s strong claim that the JTC bias has *no* role in the aetiology of delusions. Rather, I think this only shows that the JTC bias along with the anomalous experiences fall short of being a complete explanation of delusions. It is still possible that the JTC bias is one of the contributory factors in the aetiology of delusions, although it is a remaining task for cognitive-bias theorists to tell us how it is possible: such as at which stage of the development of delusions the JTC bias plays a causal role.

Can the cognitive-bias theory offer a satisfactory explanation of delusions by appealing to more than one cognitive bias? Apart from the JTC bias, in their review McLean et al. (2017) argue that delusions are also associated with the bias against disconfirmatory evidence, the bias against confirmatory evidence, and liberal acceptance. Can we expect that a combination of these cognitive biases would offer a satisfactory explanation of delusions? There are several reasons to make us less optimistic about it.

First, as of now, there are great ambiguities about how these cognitive biases are related to each other, and about “whether any of these ... biases is necessary or sufficient for the formation and maintenance of delusions” (Bronstein et al., 2019, p. 1). Though this reason allows the possibility that we might overcome this shortcoming with the development of a more sophisticated cognitive-bias theory (for a recent attempt, see Bronstein et al., 2019), the

following reasons raise some more fundamental problems with the cognitive-bias theory of delusions.

Second, regarding the evidence challenge of delusions, even though the cognitive-bias experiments show that people with delusions respond to evidence in a way different from the way people without delusions respond to the same evidence, none of them have demonstrated that the way people with delusions respond to evidence in the experiments fully captures the flagrant ways delusions flout evidence. In fact, there is good reason to doubt whether any cognitive-bias experiments can fully capture the flagrant ways delusions are formed and maintained in the light of counterevidence. The reason is that, unlike delusions, cognitive-bias-based beliefs are usually sensitive to counterevidence: when people are aware that their initial judgements are incorrect due to the influence of some cognitive biases, they usually would change their judgements. This is arguably how the participants with delusions act in the JTC bias experiments: in their review, Fine et al. (2007) observe that when participants with delusions are confronted with potentially disconfirmatory evidence to their initial judgements, they “become significantly less certain about their [initial judgements]... in comparison with nonpsychiatric, but not psychiatric, control groups” (p. 53) The striking difference between cognitive-bias-based beliefs’ being sensitive to counterevidence and delusional beliefs’ flouting of evidence suggests that, at least, cognitive biases alone would not satisfactorily explain delusions.

Third, cognitive biases are relatively domain-general, and hence have difficulties in explaining the specificity of delusions. People with delusions have displayed the JTC bias in experiments using a variety of stimuli such as beads and fish (McLean et al., 2017). It is, however, far from clear why, if the JTC bias is the main causal factor of delusions, people with delusions do not have a wide range of delusional or odd beliefs (see also Fine et al., 2007).

These problems with the cognitive-bias theory suggest that cognitive biases are unlikely to be the main causal factor in the aetiology of delusions. But it is worth re-emphasising that these problems do not demonstrate that it is impossible for any cognitive biases to play any contributory role, though it remains to be shown how exactly certain cognitive biases along with other factors can offer a satisfactory account of delusions.

3.3. The two-factor theory

3.3.1. Some clarifications

The two-factor theory may be taken as a close friend of the cognitive-bias theory in the general sense that both agree that certain reasoning abnormalities play a causal role in the aetiology of delusions. Indeed, Stone and Young (1997), who proposed an early version of the two-factor theory,⁴¹ argued that the JTC bias is a candidate for “the second factor”. More generally, Stone and Young wrote: “The second factor in our account of delusions, is, ... a reasoning *bias*.” (p. 359, emphasis in original) In a recent version of the two-factor theory, Coltheart and Davies (2021a) discuss that the bias against disconfirmatory evidence may be a candidate for the second factor. Despite that, many versions of the two-factor theory are more sceptical about taking cognitive biases as the second factor. Davies et al. (2001) write: “In our view, neither attributional biases nor data gathering biases can play the role of the second factor in the etiology of monothematic delusions.” (p. 148; see also Aimola Davies & Davies, 2009)

⁴¹ Young and colleagues proposed several early versions of the two-factor theory (Stone & Young, 1997; Young et al., 1993; Young, 1998). For critique, see Davies and Coltheart (2000), focusing on the philosophical aspects, and Breen et al. (2000b), focusing on the empirical aspects.

In any case, compared to the cognitive-bias theory, the two-factor theory has paid much closer attention to the details of the relationship between anomalous experiences and delusions, and also considered a range of reasoning abnormalities other than cognitive biases (Aimola Davies & Davies, 2009; Aimola Davies et al., 2009; Breen et al., 2000a; Coltheart, 2007, 2010; Coltheart & Davies, 2021a; Coltheart et al., 2011; Davies & Coltheart, 2000; Davies et al., 2001; Davies & Egan, 2013; Langdon & Coltheart, 2000).

The two-factor theory is sometimes interpreted as merely adding a second factor to Maher's one-factor theory of delusions. Strictly speaking, this interpretation can be misleading. This is because Maher's first factor is an anomalous experience, which may or may not be underpinned by a neuropsychological deficit. Maher (1999) writes: "[T]he model of delusion formation provided in this paper posits that the basic origin lies in the anomalous experience, regardless of how that anomaly arose." (p. 566) By contrast, the two-factor theory takes the first factor as a neuropsychological deficit which "is responsible for the belief having occurred to the person in the first place . . . : this factor determines the content of the delusional belief." (Coltheart, 2010, p. 18)

These two conceptions of the first factor can be compatible, when the neuropsychological deficit (the first factor in the two-factor theory) underpins the anomalous experience (the only factor in Maher's one-factor theory) that gives rise to delusions. But they can also come apart: Maher's one-factor theory allows the possibility that the anomalous experience is not underpinned by some neuropsychological deficit (rejecting the first factor in the two-factor theory); the two-factor theory allows the possibility that Maher's anomalous experience plays no role in the aetiology of delusions. Coltheart et al. (2010), for example, argue that the delusion, rather than Maher's anomalous experience, may be "the first delusion-relevant event of which the patient is aware" (p. 264).

The two-factor theory, as a group, also contains many versions with disagreements on important issues. Regarding the question of to what extent subjects with delusions are consciously aware of the formation of delusions, Coltheart et al.'s (2010) two-factor theory, as we have seen, argues that the formation of the Capgras delusion is not available to consciousness; McKay (2012) explicitly states that "I make no assumptions about the extent to which the generation and evaluation of candidate hypotheses are conscious." (footnote 7); Davies and Egan (2013) explore both the possibility that the Capgras delusion is the first delusion-relevant event of which the patient is aware, and the possibility that the Capgras delusion is formed via a conscious process.

Regarding the question of what the role of the second factor is in the initial adoption of the Capgras delusion and in the persistence of the Capgras delusion, Coltheart et al. (2010) and Davies and Egan (2013) agree, though for different reasons, that a second factor is not needed in the adoption stage but is needed in the persistence stage. By contrast, McKay (2012) argues that a second factor, specifically a bias towards explanatory adequacy, is needed in both the adoption and persistence stages.

More differences can be drawn among versions of the two-factor theory which are developed for different delusions, or different delusional cases in which the patients' delusions share the same content but may have various causes. Indeed, according to Aimola Davies and Davies (2009), the two-factor theory should be understood as a two-factor framework that allows "parametric variation":

There may or may not be abnormality in the first stage. If there is first-stage abnormality then the nature of the abnormality will vary from delusion to delusion and may vary from case to case of the same delusion. A first deficit may or may not give rise to an anomalous experience and the route from first deficit to belief

may lie mainly at the personal or the subpersonal level. At the personal level, an anomalous experience may have a representational content close to or far from the content of the delusion itself and the route from experience to belief may be endorsement or explanation. The personal- or subpersonal-level processes of hypothesis generation and confirmation may be subject to one or another bias within the normal range, or to frank abnormalities. (p. 315)

I concur that this is how the two-factor theory should be understood. Nonetheless, to simplify the analysis of the two-factor theory, in the following I shall focus on the version developed by Davies and Egan (2013), and set aside its differences from other versions of the two-factor theory (Aimola Davies & Davies, 2009; Aimola Davies et al., 2009; Breen et al., 2000a; Coltheart, 2007, 2010; Coltheart et al., 2011; Coltheart & Davies, 2021a; Davies & Coltheart, 2000; Davies et al., 2001; Langdon & Coltheart, 2000). In addition, I shall focus on the personal-level processes in the aetiology of delusions, and set aside issues related to the subpersonal-level processes. That is to say, I shall focus on how an anomalous experience gives rise to a delusion, and set aside questions about the aetiology of the anomalous experience, the answer to which requires a subpersonal theory.⁴² Despite these simplifications, I believe the points I shall raise are general and may be applicable to many versions and the subpersonal parts of the two-factor theory.

Davies and Egan (2013) distinguish between two ways an anomalous experience gives rise to a delusion in the adoption stage (for early discussions of this distinction, see Davies et al., 2001; Davies & Coltheart, 2000; see also Aimola Davies & Davies, 2009; Bayne & Pacherie, 2004; Langdon & Bayne, 2010; Turner & Coltheart,

⁴² Davies and Coltheart (2000) write: “[T]he way that brain damage leads to anomalies in experience is not itself a matter for folk psychology” (p. 9).

2010). According to the explanationist account, the content of the delusion is not encoded in the anomalous experience, and the delusion arises as an explanation of the anomalous experience;⁴³ according to the endorsement account, the content of the delusion is already encoded in the anomalous experience, and the delusion arises as an endorsement of the anomalous experience.⁴⁴ I will discuss these two accounts in turn, and then discuss their view on why delusions persist once they are adopted.

3.3.2. Adoption: the explanationist account

Among the three proposed anomalous experiences in Section 2.2, the experience of reduced affective responses, or the experience

⁴³ In their paper, Davies and Egan (2013) also introduce an important distinction between Bayesian inference and inference to the best explanation. Given the same evidence, Bayesian inference and inference to the best explanation may generate different results. One main reason is that while Bayesian inference only considers the probabilities of candidate hypotheses in the light of evidence, inference to the best explanation also takes into account the explanatory virtues of candidate hypotheses, including “parsimony, scope, depth, unifying disparate phenomena, and making new predictions” (p. 696)

Though the relationship between Bayesian inference and inference to the best explanation is a point of contention in the literature (see e.g. Dellsén, 2017; van Fraassen, 1989), I think Davies and Egan’s (2013) discussion raises a more general concern: that is, what kind of reasoning processes is employed by the mind? Apart from Bayesian inference and inference to the best explanation, other models of the mind include, to list a few, the Dempster–Shafer framework, the imprecise probability framework, the possibility framework, the ranking framework, and the quantum probability framework (for reviews, see Colombo et al., 2021; Genin & Huber, 2021; Halpern, 2017).

Compared with the task to model the mind at the personal level, the task to find out the best way to model the mind at the subpersonal level may be even more challenging when we consider the fact that, unlike the personal-level reasoning processes, the subpersonal-level reasoning processes are unavailable to the consciousness. Nevertheless, given that this thesis focuses on the personal-level processes, about which we have a fairly intuitive grasp of how people reason or should reason, it seems ok if we set aside the complications about these models of the mind for now. In Chapter 4, I will discuss how the general point about the shortcomings of taking reasoning abnormality as a main factor in the aetiology of delusions may be applicable to the Bayesian account of delusions as well.

⁴⁴ As regards the question of how patients have such anomalous experience in the first place, see Davies and Egan’s (2013) as well as Pacherie (2008), Wilkinson (2016), and Bongiorno (2019).

of reduced familiarity (The Affective View), and the experience that “There’s something odd about this woman.” (The Alert View) are all suitable candidates for the explanationist account’s notion of the anomalous experience in which the content of the Capgras delusion is not encoded. The question is whether the explanationist account is right that the Capgras delusion arises as an explanation of the anomalous experience.

At first sight, it seems unlikely that the person with the Capgras delusion believes that his wife is an imposter simply because the imposter hypothesis is an explanation of his anomalous experience. One reason is that there is significant evidence that the imposter hypothesis is not true: the imposter hypothesis is implausible in the light of common knowledge; it is at odds with the fact that the “imposter” looks like his wife and knows things that only his wife knows; it is also at odds with the testimonies from his friends and relatives; the subject may even have partial insights into the force of the counterevidence and the implausibility of the delusion (see Chapter 1). Another reason is that there are many more plausible hypotheses which can better explain the anomalous experience: such as the brain-damage hypothesis and the fading-love hypothesis suggested by Coltheart (2007, p. 1059).

In response, Davies and Egan (2013) offer two suggestions. First, they argue that taking the anomalous experience as such “underdescribes the way in which the [anomalous experience] ... is permeated by the sense of significance and the urgent demand for explanation and interpretation.” (p. 715; for a similar suggestion, see e.g. McKay, 2012, p. 343) Instead, “the feeling of heightened significance that suffuses the experience of the patient’s wife cries out for explanation in terms of change in the environment, not change in the patient’s brain (such as a stroke).” (p. 719)

This suggestion is, in a sense, redolent of Maher’s (1999) suggestion that “the kinds of anomalous experience that deluded

patients have appeared to be much more *intense* and *prolonged* than those that occur to the population in general” (p. 566, emphasis added). The difference is that Davies and Egan are not arguing for a one-factor theory because even the imposter hypothesis is adopted as the explanation of the anomalous experience, they argue that a second factor is needed to explain why the delusion persists.

The problem with this suggestion is, however, that the empirical evidence, i.e. the reduced autonomic responses to familiar faces, does not support that there is such an anomalous experience with the feeling of heightened significance that selectively cries out for the imposter explanation *in particular*. Nor is it clear what kind of anomalous experience can selectively cry out for the imposter explanation in particular.

Moreover, appealing to such a particularly anomalous experience would put the two-factor theory in a difficult position. On the one hand, if the two-factor theory can appeal to such a particularly anomalous experience, there is no reason in principle to prevent the one-factor theory from appealing to an even more *particularly* anomalous experience which demands the imposter explanation in both the adoption and persistence stages, and hence a second factor would not be needed at all. On the other hand, if the two-factor theory gives up appealing to such a particularly anomalous experience, then it is far from clear why the subject adopts the imposter hypothesis rather than the brain-damage hypothesis and the fading-love hypothesis.

Davies and Egan’s second suggestion is that:

[T]he cognitive processes that are engaged by a situation like that of the Capgras patient are not reflective and unbiased processes of ... inference to the most probable hypothesis all things considered, but more encapsulated and biased processes of inference to the first (most accessible) hypothesis to

predict the anomalous experience well enough (that is, with a high enough likelihood on the evidence of the anomalous experience) (p. 715)

If it is true that the adoption of the imposter hypothesis involves encapsulated and biased processes of inference, then the encapsulated and biased processes may help explain why the significant counterevidence is not properly considered before the adoption of the imposter hypothesis. In this sense, the encapsulated and biased processes may help explain why the correct hypothesis that this woman is his wife is not adopted. But without further details about the nature of the encapsulated and biased processes, it is far from clear why the imposter hypothesis, rather than the brain-damage hypothesis or the fading-love hypothesis, is adopted. What kind of encapsulated and biased processes favours the imposter hypothesis as the most accessible or best explanation? Parrott (2016) further argues that given how implausible the imposter hypothesis is, it should not even be considered as a candidate explanation of the anomalous experience in the first place, let alone as the most accessible or best explanation.

The basic problem with the explanationist account is that there is an explanatory gap between the anomalous experience as the explanandum and the delusional hypothesis as the explanans. The proposal of the encapsulated and biased processes does not seem to be able to bridge this explanatory gap. Of course, it is a fact that the delusional hypothesis is adopted. The explanation gap, however, casts doubt on the extent to which the delusional hypothesis is adopted *because* the delusional hypothesis is the most accessible or best explanation of the anomalous experience.

In an early review of the two-factor theory, Coltheart (2007) appears to acknowledge that:

The two-deficit approach [the two-factor theory] says nothing about what determines the patient's choice

between various possible [explanatory] hypotheses if all have explanatory adequacy. (p. 1059)

This is a striking acknowledgement, because if current conceptions of the two factors fall short of telling us why delusional hypotheses are favoured over other more plausible hypotheses, then either alternative conceptions of the two factors (hence a revised two-factor theory), or some additional factor (hence a multi-factor theory) must be needed.

3.3.3. Adoption: the endorsement account

According to the endorsement account, the content of the Capgras delusion is already encoded in the anomalous experience. The subject with the Capgras delusion may have an anomalous experience that “This woman looks like my wife but she is not my wife”, and the Capgras delusion is adopted as an endorsement of the anomalous experience.

In a critique of an early version of the endorsement account, Campbell (2001) writes: “Sameness of content is not alone enough to establish that the transition [from the anomalous experience to the delusional belief] is prima facie rational’ (p. 96). A reading of Campbell’s challenge is that given how anomalous the experience is, the subject should not endorse the content of the anomalous experience.

Davies and Egan’s (2013) paper contains a possible response to this challenge. They draw on Gilbert and colleagues’ finding that people’s default response to their perception is to “quickly and automatically” believe what they perceive (Gilbert, 1991, p. 107; Gilbert et al., 1990; Gilbert et al., 1993), and suggest that people with delusions may have a similar default or prepotent doxastic response to their perception. Gilbert and colleagues’ view on perception-based belief formation is known as the Spinozan view; Davies and Egan

(2013) call their view on perception-based delusion formation the near-Spinozan view. It is a *near*-Spinozan view because, as Davies and Egan emphasise, they are drawing a comparison between perception-based delusion formation and the Spinozan view on perception-based belief formation in general, but they are not committed to the stronger view that perception-based belief formation is always Spinozan (p. 708; see also Davies et al., 2001, p. 153).

The Spinozan view itself is not without controversies (for recent critique, see e.g. Hasson et al., 2005; Nadarevic & Erdfelder, 2019; Richter et al., 2009; Street & Kingstone, 2017; for defence, see e.g. Asp et al., 2020; Fazio et al., 2015; Mandelbaum, 2014; Unkelbach & Greifeneder, 2018). My present concern is not to discuss whether the Spinozan view on perception-based belief formation in general is true, but to discuss to what extent Davies and Egan's near-Spinozan view on perception-based delusion formation can help explain delusions.⁴⁵

One reason for doubting the near-Spinozan view on delusion formation is that neither Gilbert and colleagues nor the latter proponents of the Spinozan view have tested whether people with delusions act in a Spinozan way. Nor have they tested whether people in general would still act in a Spinozan way when they have experiences whose contents are as bizarre as the contents of bizarre delusions, or have experiences which face significant counterevidence as delusions do. On the face of it, given how bizarre the content of the Capgras delusion is and how much counterevidence it faces, wouldn't it be natural for the patient to inhibit the default or prepotent doxastic response and disbelieve what he perceives?

⁴⁵ It is worth re-emphasising that Davies and Egan's near-Spinozan view is only about the default or prepotent doxastic response in the initial adoption of delusions. In the literature, the phrase *the Spinozan view* is sometimes associated with other controversial commitments which should not be assumed to be held by Davies and Egan's near-Spinozan view. For a radical version of the Spinozan view on delusions, which is loaded with many other controversial commitments, see Bongiorno (2021).

In response, one might argue that the default or prepotent doxastic response is so quick that it dismisses the implausibility of the content of the anomalous experience and the related counterevidence. That is to say, at least there is a moment, no matter how short it is, the patient with the Capgras delusion may believe whatever he perceives because of the default or prepotent doxastic tendency.

The problem with this response is that even if we accept that in a split second the person adopts the content of the anomalous experience as a belief, shouldn't he immediately reject the delusion, especially when he realises how bizarre it is and how much counterevidence it faces.⁴⁶

It is worth noting that one of the views discussed by Maher may be taken to help develop the endorsement account: that is, immediate experience may have the power to outweigh reason. Maher (1999) writes:

An individual with schizophrenia with whom the writer frequently interacts has on many occasions come by rational analysis to the determined conclusion that the voices of well-known people that he hears talking to him in his room are really his own thoughts, and not the voices of the people themselves. Within minutes

⁴⁶ Recent experiments suggest that patients with schizophrenia tend to over-weigh and overcount direct experience and under-weigh information from others (Simonsen et al., 2021; Jardri et al., 2017). At first sight, this might help explain the endorsement account to explain why patients with delusions adopt the contents of their anomalous experiences.

There are, however, several problems with appealing to this specific tendency. For one thing, the experiments tested patients with schizophrenia rather than patients with delusions in particular. It remains to be seen whether patients with delusions have this specific tendency. For another, the experimental paradigm that is used to measure this tendency is more like the beads task that is used to measure the JTC bias. Therefore, my critique of the JTC bias theory of delusions is also applicable here: that is, the degree to which people with delusions over-weigh their anomalous experiences is unparalleled to the degree of the participants' tendency to over-weigh and overcount direct experience in these experiments. Therefore, this specific tendency is unlikely to be the main factor in the aetiology of delusions.

of this insight he will nevertheless begin responding directly to them, and revert to telling me what “they” are saying. The power of immediate experience over reason is quite clear. (p. 568, endnote 3)

It is not straightforward whether Maher is talking about immediate experience in general or only a particular kind of immediate experience. I think it is true that certain kinds of immediate experience may have the power to outweigh reason, and this power can help explain the adoption of delusions in the light of counterevidence. But it seems to be an implausible view that any immediate experience would have the power to outweigh reason. The immediate experience of the Müller-Lyer illusion, for example, normally does not have the power to outweigh reason such that it compels us to believe that the two lines in the Müller-Lyer illusion have different lengths. As it currently stands, however, the endorsement account seems to talk about immediate experience in general; even if we intend to develop it by focusing on a particular kind of immediate experience, it is far from clear what this particular kind of immediate experience could be, and how it can outweigh reason. (For a development of the endorsement account along this line of thought, see Section 3.2.1 of Chapter 4)

3.3.4. Persistence

Suppose that the imposter hypothesis is initially adopted as a belief because the significant counterevidence is not properly taken into consideration due to the encapsulated or biased explanatory processes or the default or prepotent doxastic tendency, the next question is why the subject does not reject the delusion later when he has all the time needed to properly consider the significant counterevidence: the delusion is implausible in light of common knowledge; it is at odds with the fact that the “imposter” looks like his

wife and knows things that only his wife knows; it is also at odds with the testimonies from his friends and relatives; the subject may even have partial insights into the force of the counterevidence and the implausibility of the delusion.

Davies and Egan (2013) argue that to properly evaluate the delusion in the light of the significant counterevidence, the subject needs to take a step back from the initial adoption of the delusional belief and re-evaluate both the imposter hypothesis and other hypotheses such as the brain-damage hypothesis and the fading-love hypothesis. These belief evaluation processes are demanding of the executive function and working memory resources. The suggestion is that subjects with delusions may have an impairment of executive function and/or working memory:⁴⁷

Impaired executive function might prevent the patient from stepping back from his initial adoption of the delusional belief; and impaired working memory might not allow the patient to work out the consequences of his prior beliefs. (p. 712)

This conception of the second factor is initially proposed by Aimola Davies and Davies (2009) on the basis of their analysis of the empirical studies of belief evaluation and their own study of anosognosia (Aimola Davies & Davies, 2009; see also Aimola Davies et al., 2009; Davies et al., in press).

I agree that this proposal can help explain why delusions are not rejected in the light of counterevidence and other more plausible hypotheses, but it seems at odds with the specificity of delusions. That is, if people with delusions suffer from an impairment of executive function and/or working memory, this impairment, as a

⁴⁷ This conception of the second factor is more specific than the standard conception of the second factor as an impairment of the belief evaluation system. For the neural basis of the second factor, see Coltheart et al. (2018).

domain-general deficit, would have prevented the subjects from properly evaluating not only delusional beliefs but beliefs in general, and consequently they would have had a wide range of delusional or odd beliefs. But this is not the case in delusions.

The more general proposal that the second factor is an impairment of the belief evaluation system faces another problem: that is, it has counter-intuitive predictions. Consider illusions. We arguably also have a tendency to believe what we perceive in the case of illusions, but we can re-evaluate whether what we perceive is true and reject the beliefs based on illusions. If people with delusions suffer from an impairment of executive function and/or working memory which prevents them from performing proper belief evaluation processes, then they would incorrigibly believe what they perceive in every illusion. Are people with delusions so vulnerable to being deluded by illusions? Davies and Coltheart (2000) write: “[W]e are very doubtful that a Capgras patient would inevitably be taken in by every illusion.” (p. 25; see also Davies et al., 2001, p. 153) And I agree. But it is not clear why the impairment of executive function and/or working memory selectively prevents the subjects from properly re-evaluating their delusions, but does not prevent them from properly re-evaluating illusion-based beliefs.

In response to the specificity challenge, Coltheart et al. (2011) offer the following suggestion:

[T]he belief evaluation system is impaired rather than abolished ... Given this, perhaps odd ideas that sporadically present themselves can be rejected by even an impaired belief evaluation system. In contrast, the idea that prompts a monothematic delusion is persistently present: every time the Capgras patient sees his wife, there will be a mismatch between the expected and the obtained autonomic response, so there is a continuous ongoing presence of the abnormal datum that is explained by the Capgras

abductive inference, and the effect of that datum can be continuously resisted only if the belief evaluation system is fully intact. (p. 289)

The key idea is that the belief evaluation system is only impaired to the extent that it cannot work properly in response to anomalous experience but can work properly in response to ordinary experience (perhaps including illusory experience). The underlying assumption is that there is a substantial difference between anomalous experience and ordinary experience, which can make the otherwise workable belief evaluation system produce delusions. The question is whether there is such a substantial difference between anomalous experience and ordinary experience.

Coltheart et al. suggest that in the Capgras delusion the substantial difference is the persistent presence of the anomalous experience. This suggestion is, once again, redolent of Maher's (1999) suggestion that "the kinds of anomalous experience that deluded patients have appeared to be much more *intense* and *prolonged* than those that occur to the population in general" (p. 566, emphasis added). Defenders of Maher's one-factor theory can agree with Coltheart et al. that there is a substantial difference between anomalous experience and ordinary experience. But, disagreeing with two-factor theorists, they hold that the substantial difference can make an *unimpaired/normal* belief evaluation system produce delusions and there is no need for a second factor. It is not clear how we can settle this disagreement between one-factor and two-factor theorists on the question of how anomalous the anomalous experience is.

In an early critique of Maher's one-factor theory, Davies et al. (2001) argue that one problem with Maher's one-factor theory is that it is unclear how to "quantify" the difference between anomalous experience and ordinary experience. It seems that the same critique is applicable to the present defence of the two-factor theory: that is,

it is unclear how to quantify, in the two-factor theory, how anomalous the anomalous experience is and how much impaired the impaired belief evaluation system is.

Without further elaborations on the difference between anomalous experience and ordinary experience and the degree of the belief evaluation impairment, we are left with, at least, three competing hypotheses: the difference can make an unimpaired/normal belief evaluation system produce delusions (Maher's one-factor theory); the difference can make a partly impaired belief evaluation system produce delusions (the two-factor theory); the difference cannot make any distinctive effect on the belief evaluation system, no matter it is impaired or not. It is far from clear how to break the impasse. Appealing to the substantial difference between anomalous experience and ordinary experience and the degree of the belief evaluation impairment, therefore, does not seem to be a promising way to defend the two-factor theory.

In sum, it is probably true that subjects with delusions do suffer from an impairment of executive function and/or working memory (Aimola Davies & Davies, 2009; Aimola Davies et al., 2009; Davies et al., in press). The question is whether the impairment is the main factor that prevents a proper evaluation of delusions. Its difficulties in addressing the specificity challenge seem to suggest that even if there is a belief evaluation impairment in the aetiology of delusions, such an impairment along with the anomalous experiences cannot be the whole story of why delusions are not properly evaluated and rejected. An alternative candidate for the second factor (hence a revised two-factor theory) or some additional factor (hence a multi-factor theory) may be needed.

So far in this section, I have argued that compared to anomalous experience which can help address the specificity challenge but have difficulties in addressing the evidence challenge,

reasoning abnormality, which is proposed to be a candidate for the second factor, can help us address the evidence challenge of delusions but have difficulties in addressing the specificity challenge. As we have seen, the two-factor theory takes into account both anomalous experience and reasoning abnormality and makes significant improvements, but it also inherits some of the shortcomings of both, and has explanatory gaps in addressing the evidence and specificity challenges.

What does this tell us? Like that the shortcomings of Maher's one-factor theory do not entail that anomalous experience is not a causal factor in the aetiology of delusions, I think the shortcomings of the cognitive-bias theory and the two-factor theory do not entail that there is no reasoning abnormality in the aetiology of delusions. Overall, the suggestion seems to be that there must be *something* in delusions, which has not been captured by the notion of anomalous experience and the notion of reasoning abnormality.

4. Something is missing

What is this something? An intriguing suggestion by Campbell is that we should not think of delusions as (normal or abnormal) responses to anomalous experiences, rather there is something closely related to delusions themselves, which cannot be fully captured by their relationship with anomalous experiences. Let us turn to Campbell's theory.

Campbell (2001) seems to agree with the endorsement account that the content of the delusion is (partly) encoded in the anomalous experience. Regarding the Capgras delusion, Campbell suggests that the anomalous experience may be that "This person is not that [remembered] person, but an impostor." (p. 96) But, according to Campbell, the Capgras delusion does not arise as a response to the anomalous experience. On the contrary, the

anomalous experience is caused by the delusion. He writes: “[T]here is some top-down loading of the perception by the delusional content.” (p. 96)⁴⁸

If Campbell is correct that the anomalous experience is not the cause of the delusion, then what is the cause of the delusion? Campbell has offered three proposals. Let us consider them in turn.

First, a delusion is a “direct result of organic malfunction” (Campbell, 2001, p. 97). The idea that an organic malfunction can be a *direct* cause of a delusional belief may strike us as radical. Campbell (2008, 2013) later develops an interventionist account of causation which offers us reason to think it is a theoretical possibility. According to the interventionist account of causation (Woodward & Hitchcock, 2003), “for X to be a cause of Y is for X to be correlated with Y under interventions on X”, if the intervention viable I meets the following conditions:

1. I causes X.
2. I does not cause Y otherwise than by X.
3. I is not correlated with any Z causally relevant to Y otherwise than via X.

⁴⁸ Campbell’s theory is often called a top-down theory of delusions, as opposed to a bottom-up theory of delusions (see Bayne & Pacherie, 2004). In the literature, there is also a hybrid view, according to which the “understanding of delusion formation is best served by considering both top-down and bottom-up mechanisms in tandem, not by prioritizing one at the expense of the other.” (Hohwy, 2004, p. 67) Another overlooked possibility is that there is no direct interaction between anomalous experiences and delusions, while they may or may not have shared causal factors: for example, it might be that the reduced autonomic activity may engender the anomalous experience and the Capgras delusion via independent routes. Nonetheless, it is important to note that the terms *top-down* and *bottom-up* are mainly concerned with the relationship between anomalous experiences and delusions. Strictly speaking, the top-down view that “there is some-top-down loading of the perception by the delusional content” is not a theory of how delusions come about but a theory of how anomalous experiences come about. The top-down view is compatible with various theories of the aetiology of delusions, a candidate for which is Campbell’s suggestion that a delusion is a “direct result of organic malfunction”. In the present section, I am mainly concerned with the question of how delusions come about; and I am not concerned with the separate question of whether there is a top-down loading after the delusion has arisen.

4. I suspends X from its usual causes. (Campbell, 2013, p. 937)

That is to say, if there is empirical evidence that certain organic malfunction is correlated with delusions under interventions on the organic function and the interventions meet the above conditions, then the organic malfunction is a cause of delusions. The problem with this proposal is, however, that so far no such empirical evidence has been found.

Second, people with delusions have lost their grasp of the meanings of the words used in their delusions. This proposal may seem attractive when we consider how difficult it is for us to understand the meanings of certain delusions. Consider the Cotard delusion, what does the person mean when she claims that she is dead? Does she really use the word “dead” in the same sense as we do? Since people normally do not use the word “dead” in the same way as people with the Cotard delusion, it seems as if the word “dead” has a different meaning for people with the Cotard delusion.

This proposal comes with a testable prediction: that is, if we examine the general knowledge possessed by people with the Cotard delusion, we should find some abnormalities in their knowledge about being alive or dead. Campbell did not consider this prediction. But it seems to me that this prediction is at odds with the empirical evidence. For example, Young and Leafhead (1996) examined the general knowledge of individuals with the Cotard delusion, and they found that, apart from claiming themselves to be dead, those individuals had no abnormality in their “general knowledge of whether other people were alive or dead.” (p. 156)

Third, it is proposed that delusions are framework propositions (Campbell, 2001; Eilan, 2001). The concept of framework propositions is put forward by Wittgenstein (1969) when he talks about a certain kind of propositions such as: “I have forebears ... every human being has them ... the earth is a body on whose surface we move and ... it

no more suddenly disappears or the like than any other solid body” and “This is my hand” (§234; §412). A negative feature of framework propositions is that they are not produced by reasoning processes on the basis of evidence. They “like the axis around which a body rotates. This axis is not fixed in the sense that anything holds it fast” (§152). A positive feature is that framework propositions act as the foundations of our belief system: other beliefs are systematically based on them and doubting a framework proposition means “toppling all other ... [related beliefs] with it” (§419).

The positive feature of framework propositions is, however, at odds with the specificity of delusions, especially monothematic delusions, for delusions are usually not systematically related to the subjects’ other beliefs (Bayne & Pacherie, 2004, p. 9; Bortolotti & Broome, 2008, pp. 832, 834-835). The negative feature of framework propositions is able to capture the sense that delusions are not based on evidence. But this appears to be a redescription of the puzzle of delusions and falls short of telling us why it is so. (For other critiques of the idea that delusions are framework propositions, see e.g. Broome, 2004; Gallagher, 2009; Klee, 2004; Thornton, 2007)

Despite these problems with thinking of delusions as the direct results of organic malfunction, as the results of the lost grasp of meaning, or as framework propositions, I think they do capture a very important aspect of delusions: that is, thinking of delusions in terms of their relationship with anomalous experiences cannot fully capture the reason why they are adopted and maintained; the missing factor may be something closely related to delusions themselves.

To see that there is something closely related to delusions themselves, which cannot be captured by their relationship with anomalous experiences, it may also be helpful to compare delusional statements with ordinary statements. Consider the Cotard delusion that “I’m dead”, the related ordinary statements are that someone is

dead and that someone is alive. When do ordinary people say that someone is dead or that someone is alive, and for what reason?

The statement that someone is dead is often used as an explanation. This is what parents may say to their innocent children when they ask why their granny lies still on the bed of the hospital and does not respond to their calls. By saying that “Granny is dead”, the parents try to explain to their children why their granny lies still on the bed of the hospital and does not respond to their calls.

Rarely, however, do we use the statement that someone is alive as an explanation, at least not in the same sense. The belief that I am alive, you are alive, she is alive, he is alive, and we are all alive is not primarily an explanation of, say, the fact that we can walk and talk.

Surely, the statement that someone is alive can sometimes be an explanation. For example, if one asks why Pinocchio can walk and talk, a reasonable answer is that “because Pinocchio is no longer a wooden puppet. He is alive”. The statement that Pinocchio is alive is an explanation of the phenomenon that Pinocchio can walk and talk. But if one asks why people can walk and talk, the answer that “because we are alive” would not strike us as a good explanation, if it is an explanation at all.

Turning to the delusional statement that “I’m dead”, the immediate question is: is it more like the ordinary statement that someone is dead, or the ordinary statement that someone is alive? The reason for the former option is obvious: the term *dead* is adopted in both the delusional statement that “I’m dead” and the ordinary statement that someone is dead, which is a product of reasoning or explanatory processes in the light of the evidence. I suspect this is a motivation for the cognitive-bias theory and the explanationist account of delusions. But the shortcomings of both in explaining delusions suggest that anomalous experiences and reasoning abnormalities, as they are currently understood, cannot be the full story of why delusions are adopted and maintained.

I think it is worth exploring that the delusional statement that “I’m dead” is more like the ordinary statement that someone is alive.⁴⁹ Like the ordinary belief that we are alive, about which there is something that cannot be fully captured by its relationship with the fact that we can walk and talk, there seems to be something about the Cotard delusion itself, which cannot be fully captured by its relationship with the person’s anomalous experience. The question is what this something is if it is not some underpinning organic malfunction, not the meanings of the words employed in the delusion, not the status of being a framework proposition.

How can we find this something? When Aimola Davies and Davies (2009) summarise the challenges faced by theories of delusions, they write that one of the fundamental challenges is “that we do not have an articulated, still less a computationally implemented, model of normal believing.” (p. 289) This suggests that perhaps we should start looking for this something in a model of normal believing. And this is what I am going to do in Chapter 3, in which I will develop a new framework of normal believing. In Chapter 4, I will return to delusions and explain how this new framework can help advance our understanding of delusions.

⁴⁹ Similar points can be made about other delusions. Consider the ordinary statements related to the Capgras delusion: that she is my wife and that she is a stranger. When an ordinary person sees a stranger and believes that she is a stranger, the stranger belief might be an explanation of his experience of seeing the stranger. But when an ordinary person sees his wife and believes that she is his wife, the wife belief does not seem to primarily function as an explanation of his experience of seeing his wife. The cognitive-bias theory and the explanationist account may hold that the Capgras delusion is more like ordinary people’s stranger belief. By contrast, I suggest that the Capgras delusion may be more like ordinary people’s wife belief.

Chapter 3.

Believing what we clearly perceive

What the experience of reality is in itself can hardly be deduced nor can we compare it as a phenomenon with other related phenomena. We have to regard it as a primary phenomenon which can be conveyed only indirectly. Our attention gets drawn to it because it can be disturbed pathologically and so we appreciate that it exists.

(Jaspers, 1913/1997, p. 94)

[T]he nature of my mind is such that I cannot but assent to these things, at least so long as I clearly perceive them.

(Descartes, 1984, p. 45)

1. Introduction

Current theories of delusions typically appeal to anomalous experiences and/or reasoning abnormalities to explain delusions. If we take it that the anomalous experiences are the subjects' evidence in favour of their delusions and the reasoning abnormalities are implicated in the subjects' evidence-evaluation ability, then current theories appear to think of delusions as a form of evidence-based beliefs. The shortcomings of current theories in addressing the evidence and specificity challenges, however, suggest that thinking of delusions as a form of evidence-based beliefs falls short of fully capturing the nature of delusions. Now the question is: what is the alternative way that can help us get a better grip on delusions? In light of Aimola Davies and Davies' (2009) suggestion that we need a better model of normal believing to understand delusions, the question is:

what is the alternative kind of normal believing that can help us get a better grip on delusions?

We may find inspiration in the self-reports of people with delusions and the reports of their close relatives. In these reports, there seems to be a recurrent indication that there is some important resemblance between delusions and certain mundane beliefs and mathematical ideas. For example, talking of her own delusions on Twitter, Mithen writes: “They feel like you know something, like you know the sky is blue.”⁵⁰ Reporting on her husband’s delusions, Kane writes: “He ‘knows’ his delusions are true the same way I know I have 5 fingers on each hand.”⁵¹ Talking of his own delusions, Nash says: “[They] came to me the same way that my mathematical ideas did.” (quoted in Nasar, 1998, p. 11)

What do these brief remarks tell us about delusions by comparing delusions to these non-delusional beliefs? It depends on how we understand the nature of these non-delusional beliefs, and what the resemblance between delusions and these non-delusional beliefs is.

One view might be that these remarks tell us nothing but that, for people with delusions, their delusions are as real as some mundane facts. This way of thinking of delusions is well-known in the literature. And it helps us to have a grasp on what it is like to have delusions, and how delusions may lead to other mental states and actions. However, it does not tell us a lot about the aetiology of delusions.

A more informative view may be that these remarks are not only comparing delusions to non-delusional beliefs but also comparing the abnormal experiences that give rise to delusions to the ordinary experiences that give rise to non-delusional beliefs.

⁵⁰ <https://twitter.com/mithenzoe/status/1364426104564449281?s=11>

⁵¹ https://twitter.com/syl_kane/status/1364461903767097344?s=11

According to this view, these remarks suggest that the abnormal experiences are as real as the ordinary experiences. Jaspers (1913/1997) seemed to have discussed a version of this view, when he considered the possibility that it is the abnormal experiences with “the character of reality” that give rise to delusions.

All primary experience of delusion is an experience of meaning, and simple, ‘one-stage’ delusional notions do not exist. For example, a patient suddenly has the notion that a fire has broken out in a far-away town (Swedenborg). This surely happens only through the meaning he draws from inner visions that crowd in on him with the character of reality? (p. 103)

Jaspers’ view that the character of reality is an important factor in the aetiology of delusions is quite intuitive. It is, however, not easy to articulate what the character of reality is. One candidate conception of the character of reality was critically discussed by Jaspers. Introducing this conception, Jaspers wrote:

Let us now try to imagine what the psychological significance is of this delusional experience of reality in which the environment offers a world of new meanings. All thinking is a thinking about meanings. If the meaning is perceived directly with the senses, if it is *directly present* in imagination and memory, the meaning has the character of reality. (p. 99, emphasis added)

According to this conception, the character of reality refers to something that supervenes on, if it is not identical to, the fact that the content of (delusional) beliefs is directly present in the subjects’ experiences. Can this conception of the character of reality explain why delusions are adopted and maintained in the face of

counterevidence? Jaspers' answer appeared to be negative. He wrote:

Immediate experience of reality survives only if it can fit into the frame of what is socially valid or can be critically tested. Experience of reality leads us to judgments of reality. Individual experience can always be corrected... (p. 104)

If it is true that this kind of experience can always be corrected, especially when it is critically tested in the face of counterevidence, then it falls short of explaining why delusions resist correction.⁵²

Importantly, the shortcoming of this specific conception of the character of reality does not disprove Jaspers' general idea that there is a shared phenomenal character between the abnormal experiences that give rise to delusions and ordinary experiences that give rise to non-delusional beliefs, and this shared phenomenal character is related to the phenomenal character of our experience of reality. In other words, even though the specific conception critically discussed by Jaspers is not the right option, it is still possible that there may be a different conception of the shared phenomenal character that can help explain delusions. At least, it can be a working hypothesis that

⁵² Current phenomenologist theories of delusions have developed various ways to understand the character of reality (see e.g. Feysaerts et al., 2021a; Ratcliffe, 2008a; Sass & Pienkos, 2013). This thesis, however, may not be the best place to engage with the phenomenological literature for several reasons. First, phenomenologist theories of delusions are less concerned with monothematic delusions, which are the primary focus of this thesis. Second, phenomenologist theories of delusions rarely talk about the relationship between Cartesian clarity and belief, which is the primary concern of this chapter. Third, phenomenologist theories of delusions are developed following a tradition of thoughts from Husserl and Heidegger. It is relatively independent of the methodology adopted by cognitive theories of delusions and is also relatively independent of the analytic literature on the relationship between experience and belief, which I shall focus on. Of course, none of these suggests against the possibility of an approach to delusions that integrates current phenomenological theories and cognitive theories (see e.g. Nelson & Sass, 2017; Humpston et al., 2019; Gallagher, 2018; Roessler, 2013), and perhaps also integrates them with the dual-force framework of delusions that this thesis proposes. In this thesis, however, I have a relatively modest aim to focus on developing the dual-force framework of delusions.

there is a distinctive kind of phenomenal character such that the subject of the experience with this phenomenal character cannot but believe what they experience, even in the face of counterevidence.

What kind of experience could it be such that the subjects cannot but believe what they experience? For philosophers, this question is likely to be reminiscent of Descartes' famous remark that "[T]he nature of my mind is such that I cannot but assent to these things, at least so long as I clearly perceive them." (Descartes, 1984, p. 45) This points to a line of research, according to which we should compare delusional cases to the non-delusional cases in which the subjects cannot but believe what they clearly perceive. We cannot but believe that the sky is blue when we clearly perceive that the sky is blue; we cannot but believe that we have 5 fingers on each hand when we clearly perceive that we have 5 fingers on each hand; and we cannot but believe that 1 plus 2 is 3 when we clearly perceive that 1 plus 2 is 3.

Surely, people with delusions do not clearly perceive their delusions (e.g. the person with the Capgras delusion does not clearly perceive an imposter), if we take perception verbs as factive, or, in Ryle's (1949/2009) terms, as "record[ing] observational successes" (p. 201).⁵³ And the resemblance between delusions and believing what we clearly perceive cannot be that both involve factive clear perception.

However, it seems plausible that there is a phenomenal character that is present in both cases. More specifically, the phenomenal character may be the phenomenal clarity of the

⁵³ Most delusions are false, and the falsehood entails that what the subjects believe to be the case is in fact not the case. Some delusions can *incidentally* be true. A person with the Othello syndrome may have the delusion that his partner is unfaithful, and this delusion may unfortunately be true, or may sometimes become true due to the person's delusion (Fulford, 1989, p. 204). But, even in such cases, the person does not clearly perceive the delusion because the truth of the delusion itself does not entail that it is a successful perception.

experience in which it seems clear to the subject that p.⁵⁴ Following this line of thought, the key to understanding delusions is to understand phenomenal clarity and its relationship with belief.

Here is the plan for this chapter. Section 2 will provide more *prima facie* motivations for considering phenomenal clarity as a factor in the aetiology of delusions, and outline the work that needs to be done for taking it seriously. Section 3 will elucidate the nature of phenomenal clarity. I will call experiences with this notion of phenomenal clarity *clear seeming experiences, or clear experiences* for short, in which something clearly seems to be so to the subject.⁵⁵ Section 4 will begin by discussing the reasons for accepting the possibility that clear experience can compel assent. Given that it is possible, I will move to discuss various explanations for how it is possible. I will argue that it is possible because of the asymmetry between the way the justificatory force and causal force of clear experience come apart and the way the justificatory force and causal force of evidence come apart. Section 5 will summarise the basic points of the previous discussion, and propose a dual-force framework of believing, according to which many beliefs can be understood as the results of the interaction between the phenomenal force of clear experience and the evidential force of evidence. Section 6 will illustrate the explanatory power of the dual-force framework by discussing how it can help advance our understanding of akratic beliefs, which are another exemplary kind of beliefs that flout evidence. I will return to delusions and discuss how the dual-force

⁵⁴ I use the term *proposition* in the generic sense. For example, when I say that a person believes that p or perceives that p, I mean that p is the content of that person's mental episode, without a particular metaphysical commitment to the nature of proposition (for a review of various metaphysical views on the nature of proposition, see e.g. McGrath & Frank, 2020). By perception, I mean conscious perception and I set aside issues related to unconscious perception (see e.g. Phillips & Block, 2016; Peters et al., 2017).

⁵⁵ I thank Johannes Roessler and Thomas Crowther for helping me locate the expressions related to phenomenal clarity.

framework can help advance our understanding of delusions in the next chapter.

2. Phenomenal clarity compels assent: first pass

The discussion of phenomenal clarity dates back to Descartes.⁵⁶ Very roughly, according to the phenomenal reading of Descartes (Patterson, 2008; Paul, 2020),⁵⁷ the notion of phenomenal clarity refers to the distinctive phenomenal character of clear experience in which it seems clear to the subject that *p*.⁵⁸ Phenomenal clarity can be present in many sorts of experiences, including intellectual experiences, such as when we consider the proposition that $1 + 1 = 2$, and sensory experiences, such as when we look at this page (Paul, 2020, p. 10). Phenomenal clarity “come[s] in degrees” (Paul, 2020, p. 4). Phenomenal clarity can compel assent, and its compelling force also comes in degrees. The most famous example in which phenomenal clarity is in its strongest form is the perception of the *cogito*, with the most compelling force. Descartes (1984) wrote: “[T]he nature of my mind is such that I *cannot but assent* to these things, at least so long as [to a certain degree] I *clearly* perceive them.” (p. 45, emphasis added; see also p. 27; p. 48)⁵⁹

⁵⁶ The discussion of the experience of reality can be traced back to much older times, perhaps to the very beginning of philosophy. For the influence of other philosophers on Descartes, see e.g. Cottingham (1992, 1994, 1998).

⁵⁷ The phenomenal reading is often proposed as an alternative to what has been called the intentional reading of Descartes’ clear and distinct perception. For reviews, see Patterson (2008) and Paul (2020).

⁵⁸ The present discussion is neutral about the nature of experience: it is not committed to, for example, representationalism or naive realism (Crane & French, 2021).

⁵⁹ Descartes thought that ordinary people often fail to distinguish a *genuinely* clear and distinct perception from a *seemingly* clear and distinct perception, in which case the subject *merely thinks* they have a genuinely clear and distinct perception. Descartes (1984) wrote: “[T]here are few who correctly distinguish between what they in fact [clearly and distinctly] perceive and what they think they [clearly and distinctly] perceive, for few are accustomed to clear and distinct perceptions” (p. 348). This was why Descartes aimed to “teach” us about the genuinely clear and

There are many important questions revolving around Cartesian clarity: for example, how can it help us deal with scepticism? Is there a vicious circle in Descartes' argument that whatever he perceives clearly and distinctly is true? And is the phenomenal reading the best way to interpret Descartes' work? These questions are not my immediate concerns. Here, I am interested in the idea that phenomenal clarity compels assent, particularly the possibility that phenomenal clarity may compel assent to the extent that the subject cannot but assent. This is because it may offer a potential explanation of the baffling phenomenon in which it seems that people cannot but assent to something in the face of counterevidence, such as in many cases of delusions.

A similar point seems to have been raised by Hobbes who agreed with Descartes that clear experience compels assent, but argued that the compelling force of clear experience may also be a factor in the aetiology of obstinate beliefs. Hobbes wrote:

[A]nyone who is free from doubt claims he has such 'great light' [i.e. great clarity] and has no less strong a propensity of the will to affirm what he has no doubt about than someone who possesses real knowledge. Hence this 'light' can explain why someone obstinately defends or holds on to a given opinion, but it cannot explain his knowledge of its truth. ("Thirteenth objection", in Descartes, 1984, p. 134)

distinct perception in his *Meditations* (Descartes, 1991, p. 165). It is, however, not straightforward whether Descartes thought that a seemingly clear and distinct perception can also compel assent. Very roughly, the notion of clear perception might be used to refer to a number of different experiences: (1.) Descartes' *genuinely clear and distinct perception* both compels assent and is a guarantee that the perceptual content is true; (2.) what I call *clear seeming experience* or *clear experience* may compel assent but is not a guarantee that the content is true; (3.) *seemingly clear and distinct perception* is a kind of experience in which the subject merely thinks that it is a genuinely clear and distinct perception, but seemingly clear and distinct perception is no guarantee that the content is true, and it might not compel assent.

Hobbes did not offer details about the kind of doxastic phenomena he had in mind. But delusions seem to be obvious candidates. On the face of it, it appears to be able to capture at least part of what is going on in some delusions: one of the reasons why a person with the Capgras delusion cannot but believe that his wife is an imposter in the face of the counterevidence may be that he has a clear experience in which it seems very clear to him that his wife is an imposter, and this clear experience compels his assent.

The more general view that a certain kind of phenomenal character of the abnormal experiences is a contributory factor in the aetiology of delusions is not new. As we have seen, Jaspers (1913/1997) had considered a version of this view and critically discussed the extent to which a specific conception of the character of reality can help us understand delusions. Most recently, Roessler (2013) argues that, in the case of thought insertion (the delusional belief that some of the subjects' thoughts are inserted or put into their mind), the subject may be in a distinctive kind of altered state of consciousness which "leads them to take at face value" their delusional idea. (p. 670)⁶⁰ Following this line of thought, the notion of phenomenal clarity can be taken as a candidate for the phenomenal character of the abnormal experiences in delusions. To my best

⁶⁰ Roessler's theory is one that integrates both phenomenological and cognitive approaches to thought insertion. According to the phenomenological theories, people with thought insertion are in an altered state of consciousness (Parnas et al., 2002; Parnas & Sass, 2001). According to the cognitive theories, the delusion is an explanation of the subjects' anomalous experience that they do not feel as "being the agent of" their thoughts (Stephens & Graham, 1994, p. 1; Campbell, 1999). Roessler (2013) argues that the phenomenological theories do not tell us about the "patients' reasons for holding that [delusional] belief" while the cognitive theories do not fully address the question of why the patients take the delusional explanation at face value given how bizarre and unusual it is. He suggests that there is a way these two groups of theories can "work in tandem" (p. 668). More specifically, he proposes that the patients may be in a distinctive kind of altered state of consciousness which "leads them to take at face value the idea of an utterly unusual, quasi-mechanical explanation of their thinking, previously invoked within the scope of 'it seems to me as if.'" (p. 670; see also Roessler, 2015; forthcoming).

knowledge, this thesis is the first attempt to explore the possibility that phenomenal clarity may help us explain why people cannot but assent to something in the face of counterevidence, especially in the case of monothematic delusions.

For this possibility to become a serious contender, it is of paramount importance to be explicit about what the phenomenal character is when we talk about phenomenal clarity. At this point, we only very roughly take phenomenal clarity as the phenomenal character of clear experience in which it seems clear to the subject that *p*. It needs significant qualifications so that we can distinguish it from related phenomena. I will do this in Section 3.

It is also of paramount importance to be explicit about why we should accept that it is possible that a person cannot but believe what seems very clear to them, and to explain how it is possible. At this point, we only have a very vague idea that it is implicit in Descartes' work and explicitly pointed out by Hobbes. We do not have an account for what exactly is going on here when we say that phenomenal clarity or clear experience compels assent. Is it the fact that the subject is in primitive contact with reality? Is it the fact that the subject *thinks* that they know that *p* in virtue of the phenomenal clarity of the clear experience that *p*? Is it because clear experience offers *prima facie* justification? Are there other explanations? I will discuss these questions in Section 4.

3. Phenomenal clarity is a distinctive phenomenon

Clear experiences are everywhere. Right now, I have a clear visual experience in which it seems clear to me that a squirrel is playing in the yard, and a clear auditory experience in which it seems clear to me that my neighbour is practising her French horn. It is, however, not an easy task to articulate what the distinctive phenomenal clarity of clear experience is. Paul (2020) writes:

Like phenomenal qualities in general, clarity is (epistemically) primitive in the sense that we cannot come to understand what clarity is by analyzing it or defining it in terms of other properties. Instead, we come to understand what clarity is by reflecting on examples, within our own experience, of clarity itself. (p. 2)⁶¹

I agree that phenomenal clarity is primitive in the sense that it cannot be defined in terms of other phenomenal qualities such as redness. I also agree that reflecting on examples is a helpful way to understand what phenomenal clarity is. But, apart from what appears to be a hopeless attempt to search for a reductive definition of phenomenal clarity, and apart from merely reflecting on cases in which phenomenal clarity is present, we may get a better grip on what phenomenal clarity is by considering what phenomenal clarity is not. In the rest of this section, I shall try to draw an outline of what phenomenal clarity is by distinguishing it from some phenomena that might be confused with phenomenal clarity to various extents.

3.1. Vividness

The notion of phenomenal clarity is sometimes confused with the notion of vividness. When we say that we have a vivid memory of an event, we usually mean that the memory is vivid to the extent as if we are seeing the event. This use of vividness indicates that vividness is the phenomenal character that is often present in its full-blooded form in perception but in a lesser form in memories as well as in dreams and imaginings. Vividness is a key notion in Hume's work (1739/2007), in which Hume used it to distinguish impressions from

⁶¹ Descartes (1984), by contrast, wrote: "[Clear perception] is something that it is easier to learn by examples than by rules," implying that there might be certain rules for identifying clear perception (p. 116). But it is not obvious what these rules are.

ideas. It turns out to be extremely difficult, however, to pin down what exactly the notion of vividness refers to. Kind (2017), for example, argues that, with respect to imaginings, “the notion of vividness ultimately proves to be so problematic as to be philosophically untenable.” (p. 33)⁶²

For our purposes, it is worth noting that there is a use of the notion of vividness, according to which the phenomenal character it refers to is similar to the phenomenal character the notion of phenomenal clarity refers to. Consider the following quotation in which Hume (1739/2007) talked about the vividness of idea:

An idea assented to feels different from a fictitious idea, that the fancy alone presents to us: And this different feeling I endeavour to explain by calling it a superior force, or vivacity, or solidity, or firmness, or steadiness. This variety of terms, which may seem so unphilosophical, is intended only to express that act of the mind, which renders realities more present to us than fictions, causes them to weigh more in the thought, and gives them a superior influence on the passions and imagination. (Appendix, p. 69)

Here Hume mentioned at least two characteristics of vividness that are similar to those of phenomenal clarity. One is that the conception of vividness as the way our mind “renders realities more present” (assuming that it is in our experiences that our mind renders realities more present) is similar to our conception of phenomenal clarity as the phenomenal character of our experiences of realities. The other is that the conception of vividness as something that “causes them to weigh more in the thought” is similar to our conception of phenomenal clarity as the phenomenal character of

⁶² For critique, see e.g. Tooming and Miyazono (2020).

experiences that compels assent. Indeed, Descartes also used the term *vivid* to talk about sensory clarity (Paul, 2020, p. 4).⁶³

Right after the quotation, Hume continued: “Provided we agree about the thing, ‘tis needless to dispute about the terms.” This might be true with respect to Hume’s main philosophical concerns. But, for our discussion, it can be quite misleading to use phenomenal clarity and phenomenal vividness interchangeably. The immediate reason is that the term *vivid* has a basic sense in which it refers to the brightness of colours.⁶⁴ This is not something we are trying to talk about when we talk about the phenomenal character of the experiences of realities.

To see the point, let us consider the case in which there are two shirts that are exactly the same except that one is vivid pink and the other is pale pink; and suppose that our visual experience of the vivid pink shirt is more vivid than our visual experience of the pale pink shirt; now the question is whether, in our experiences, the vivid pink shirt is presented as more real than the pale pink shirt. I think it is doubtful that one is more real than the other. (There is no doubt that many people like objects with bright colours. But it is doubtful that it is because objects with bright colours seem more real to them.)

Another reason for not using phenomenal clarity and vividness interchangeably is that, whereas it is relatively natural, at least for proponents of the phenomenal reading of Descartes, to say that our intellectual experience has phenomenal clarity, it is quite odd to say that our intellectual experience has vividness. Put it another way, we would have an immediate understanding of what someone means

⁶³ There is an important difference between Hume’s conception of vividness and our conception of phenomenal clarity. Phenomenal clarity is the phenomenal character of clear experience in which it seems clear to the subject that *p*. It is not the phenomenal character of the mental state of believing that *p*, if it can have a phenomenal character. For Hume, however, vividness is also something beliefs can have. He (1739/2007) wrote: “Here we must not be contented with saying, that the vividness of the idea produces the belief: We must maintain that they are individually the same.” (p. 80)

⁶⁴ <https://dictionary.cambridge.org/dictionary/english/vivid>

when they tell us that they see a mathematical equation clearly or obscurely, but we probably would struggle in understanding them if they tell us that they perceive a mathematical equation vividly or non-vividly (such expressions, if exist, may make us wonder whether they are synaesthetes).

In brief, compared with phenomenal clarity that, in both sensory and intellectual cases, refers to the phenomenal character of clear experience in which it seems to the subject that *p*, the notion of vividness may help us identify phenomenal clarity in sensory cases but it is at the price of mixing something which is not phenomenal clarity, e.g. the brightness of colours; and in intellectual cases it does not seem to be suitable for helping us identify phenomenal clarity.

3.2. Sensory information

The term *clear* is sometimes used to talk about the amount of sensory information provided by experience rather than the phenomenal clarity of experience. To see the difference, let us consider the following case:

TWO PIEBALD HORSES: Kate is standing in Port Meadow on a sunny day. A few steps in front of her there is a piebald horse. While looking at the piebald horse, Kate also sees that far away from her there is another piebald animal, but she cannot really tell whether it is another piebald horse or a dairy cow. Unbeknown to her, it is indeed another piebald horse.

There is a use of the term *clear*, according to which Kate's visual experience of the nearby piebald horse is much *clearer* than her visual experience of the faraway piebald horse. What exactly does this mean? A plausible interpretation, I think, is that Kate's visual experience of the nearby piebald horse provides her with *more visual information* than her visual experience of the faraway piebald horse:

the former provides her with the visual information about a piebald horse while the latter provides her with the visual information about a piebald animal.

This sense of being clear is different from our notion of phenomenal clarity as the phenomenal character of clear experience in which it seems clear to the subject that *p*. Though Kate's visual experience of the nearby piebald horse and her visual experience of the faraway piebald animal provide her with different amounts of visual information, it seems that both experiences have the same degree of phenomenal clarity. At least, there is no obvious reason that there is a distinction between the degree to which it seems to Kate that there is a piebald horse nearby and the degree to which it seems to her that there is a piebald animal far away. In other words, there is no obvious reason that in Kate's experiences it is more real that there is a piebald horse nearby than that there is a piebald animal far away.

The distinction between the sense in which the term *clear* is used to talk about the amount of sensory information provided by experience and our notion of phenomenal clarity can help us distinguish phenomenal clarity from the phenomenal character that concerns some phenomenal dogmatists, who may appear to talk about clear experience but are actually talking about the amount of sensory information provided by experience. Consider the following example which Koksvik (2011) uses to illustrate what he means by clear experience:

BLIZZARD: Ann is standing stationary on a flat, snow-covered plain in a blizzard. The wind is whipping snow around in all directions, and no features of the landscape are visible. Ann can barely see her own knees, and she cannot see the tips of her skis.

Someone approaches very slowly from the direction in which Ann is looking. At first she is completely unable to distinguish the approaching person from

patterns randomly forming and dissipating in the snow. As the person approaches, Ann's perceptual experience changes, the human figure gradually appears more and more clearly. (p. 187)

In line with Koksvik's description, Berghofer (2020b) writes: "The closer the person in front of Ann gets, the more clearly this person appears to Ann" (p. 165). What exactly is it that has changed or becomes "clearer" in Ann's perceptual experience, as the person approaches? Neither Koksvik nor Berghofer provides an answer. They seem to assume that the answer is obvious and move on to discuss whether, as the person approaches, Ann's perceptual experience pushes her towards believing that there is a person in front of her. Here I am concerned with what has changed in Ann's perceptual experience. A plausible answer, I think, is that what has changed is the amount of visual information provided by Ann's perceptual experience. As the person approaches, Ann gets more and more visual information about the person approaching her.

This phenomenon is, however, not what we are trying to talk about when we talk about phenomenal clarity. Like the case of TWO PIEBALD HORSES, there is no obvious reason to think that the phenomenal clarity of Ann's perceptual experience changes as the person approaches. Compare the moment at which Ann only recognises an object moving towards her and the moment at which Ann recognises that the object is Cecily: there is no obvious reason to think that at the former moment the degree to which it seems clear to Ann that there is an object moving towards her differs from the degree to which it seems clear to Ann that it is Cecily at the latter moment. That is to say, as the person approaches, the phenomenal clarity of Ann's experience may remain the same.⁶⁵

⁶⁵ As the person approaches, Ann may have a series of intellectual experiences when she considers whether it is Cecily: when she only recognises a moving object, it is obscure to her that it is Cecily; when she recognises that it is a person, depending on her background beliefs, it may be somewhat clear to her that it could

3.3. Truth

The phenomenal clarity of a clear experience in which it seems clear to us that p may be a relatively reliable marker that p is true. But, as a character of subjective experience, phenomenal clarity is certainly no guarantee that p is true.

Regarding false content, it is a familiar phenomenon that illusory or hallucinatory experiences can have phenomenal clarity. For example, in the Müller-Lyer illusion it is false that the two lines have different lengths, but it seems clear to us that the two lines have different lengths.

An experience with a false content may also have phenomenal clarity due to some neuropsychological deficit or impairment. Descartes also considered this possibility. He (1984) wrote:

For we have often noted that error can be detected in the senses, as when ... someone with jaundice sees snow as yellow; for when he sees it as yellow he sees it just as clearly and distinctly as we do when we see it as white. (p. 104)

Regarding true content, it is also quite common that a true proposition may seem obscure to us, though it may become clear to us with the help of the reasons for that proposition. Consider the following example discussed by Paul (2020, p. 11):

- a. The sum of the numbers 1, 2, and 3 is equal to their product.

According to Paul, (a.) may seem obscure to us at first.

Then consider the following proof:

be Cecily; when she finally recognises that it is Cecily, it is absolutely clear that it is Cecily. Here the terms *obscure* and *clear* are about the phenomenal clarity of Ann's intellectual experiences, which increases as the person approaches. But Koksvik and Berghofer do not seem to talk about intellectual experiences.

$$1 + 2 + 3 = 6$$

$$1 \times 2 \times 3 = 6$$

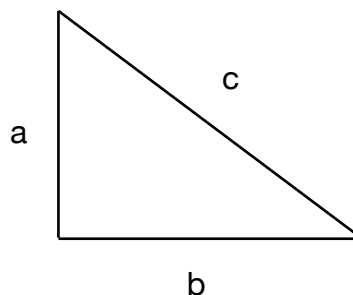
Therefore, $1 + 2 + 3 = 1 \times 2 \times 3$

It is fairly intuitive that the consideration of the proof can make a difference in the way we perceive (a.). What is the difference? According to Paul, the difference is that, with the help of the proof, our perception of (a.) now has the phenomenal clarity whereby (a.) strikes us as true.

3.4. Judgement

Phenomenal clarity and judgement are two distinct notions. To say that S has a clear experience in which it seems to S that p is not to say that S judges that p. The obvious reason is that the fact that it seems clear to S that p is only one of many possible reasons on the basis of which S may make a judgement about p. When we know that the Müller-Lyer illusion provides us with a false content, we usually would not make a judgement that its content is true, even though the illusion maintains a certain degree of phenomenal clarity.

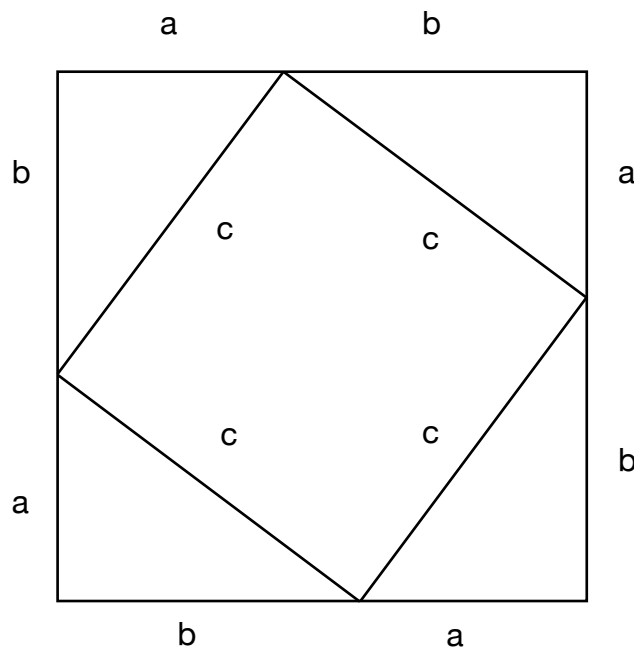
Another difference between phenomenal clarity and judgement is that phenomenal clarity can vary while S's judgement remains the same. To see the point, let us consider the Pythagorean theorem:



$$a^2 + b^2 = c^2$$

Suppose that we know the Pythagorean theorem is true but forget the reason why we believe that it is true. In this case, it is intuitive that even though we forget the reason why we believe that it is true, we still judge that the Pythagorean theorem is true.

Now let us consider the following proof of the Pythagorean theorem:



The same four triangles can form a big square. The area of the big square, $(a + b)^2$, equals the combination of the areas of the four triangles, $4 \times \frac{ab}{2}$, and the small square, c^2 :

$$(a + b)^2 = 4 \times \frac{ab}{2} + c^2$$

$$a^2 + b^2 + 2ab = 2ab + c^2$$

$$a^2 + b^2 = c^2$$

It is fairly intuitive that this proof makes a difference in how we perceive the theorem. What is the difference? The difference cannot be that with the help of the proof, we now judge that the theorem is true, because it is the same judgement before the consideration of the proof. The difference, I think, is that with the help of the proof, the way we perceive the theorem gains a degree of phenomenal clarity: it becomes clear that the theorem is true.

Similar arguments can be developed regarding other sorts of experiences. Suppose that you know that Moriarty committed the crime, but you do not know anything about the evidence discovered by Holmes: perhaps you know that Moriarty committed the crime because you read the headline of a trustworthy newspaper. In this case, it is intuitive that you judge that Moriarty committed the crime.

Now suppose that you become more interested in the case, you carefully review the evidence discovered by Holmes, and it is reaffirmed that Moriarty committed the crime. The review of the evidence, I think, would make a difference in how you perceive that Moriarty committed the crime. What is the difference? It cannot be that with the help of the evidence, you judge that Moriarty committed the crime because it is the same judgement before your review of the evidence. The difference, I think, is that with the help of the evidence, the way you perceive the proposition gains a degree of phenomenal clarity: it becomes clear to you that Moriarty committed the crime.

In both the case of the Pythagorean theorem and the case of Moriarty, I have appealed to the idea that the consideration of evidence/proof may make a difference in the phenomenal clarity of our experiences. And I have assumed that this idea is intuitive. I imagine that it might attract controversies. But it is important to note that it is not an ad hoc hypothesis for the convenience of the argument but is arguably manifested in our ordinary language.

When we talk about why we believe that p , in English we may say that we believe that p in the *light* of the evidence that p , or that the evidence casts *light* on that p . What do these light locutions mean? An evidential reading is that the light is a metaphor for the evidential relationship between evidence and belief. In the case of Moriarty, if we believe that Moriarty committed the crime in the light of the evidence discovered by Holmes, the light can be read as a metaphor for the evidential relationship between the evidence discovered by Holmes and the belief that Moriarty committed the crime. Apart from

the evidential reading, I think there may be a phenomenal reading according to which the light locutions suggest that the way we perceive *p* gains a degree of phenomenal clarity. Compare the following two sentences:

- i. The lamp casts light on the desk.
- ii. The evidence discovered by Holmes casts light on who committed the crime.

On the face of it, (i.) and (ii.) share a similar grammatical structure. This suggests that there may be a sense in which, like the desk which becomes clearer to us with the help of the lamp, the fact that Moriarty committed the crime also becomes clearer to us with the help of the evidence (For a similar surface-grammar analysis of pain locutions in English and Mandarin, see Aydede, 2019; Liu & Klein, 2020; Nie, 2021).

3.5. Psychological confidence

Talking about the phenomenal clarity of a clear experience in which it seems clear to the subject that *p* is sometimes confused with talking about the subject's confidence that *p*. Phenomenal clarity is a source of, but not identical to, psychological confidence. Other things being equal, the clearer it seems to a person that *p*, the more confident the person becomes that *p*. But phenomenal clarity and psychological confidence are dissociable: in an illusory experience it can seem clear to us that *p*, but if we know that it is an illusion, we will not feel confident that *p*; we can also feel confident that *p* when we fail to clearly see that *p*. For example, we are usually quite confident in the truth of scientific theories we read in reputable journals, but rarely can we clearly see that a scientific theory is true, especially when it is beyond our expertise.

3.6. Asserting that p

In the literature on phenomenal dogmatism (also called phenomenal conservatism) whose primary concern is perceptual justification (for early expositions, see Huemer, 2001; Pryor, 2000, 2004; for reviews, see Moretti, 2020; Huemer, 2019; Tucker, 2013b), it is argued that our experiences can have a distinctive assertiveness phenomenology: such experiences “assertively represent propositions” (Siegel & Silins, 2015, p. 793).⁶⁶ What is asserted by an experience with the assertiveness phenomenology is the truth of the content of that experience. Is the notion of the assertiveness phenomenology the same as our notion of phenomenal clarity?

There are several reasons why one might take them as the same. First, both arguably involve the fact that the content of the experience is presented as true. Second, both are gradable. Third, the examples that phenomenal dogmatists use to elucidate the assertiveness phenomenology overlap with the example we use to elucidate phenomenal clarity. These points are evident in the following quotation by Tucker (2010):

The phenomenological character of assertiveness comes in degrees. Some seemings [i.e. it-seems-to-

⁶⁶ This specific phenomenology that interests phenomenal dogmatists has been called by various names with different emphases, such as the feel of truth (Tolhurst, 1998, p. 298), phenomenal force (Pryor, 2000, p. 547), assertoric force (Chasid & Weksler, 2020, p. 733; Heck, 2000, p. 508; Siegel & Silins, 2015, p. 792), forcefulness (Huemer, 2001, p. xi), phenomenology of objectivity (Church, 2013, Chapter 1 and Section 4 of Chapter 4; Koksvik, 2011, p. 209), vivacity (Matthen, 2005, p. 306), phenomenology of pushiness (Koksvik, 2011, p. 209; 260), presentational phenomenology (Chudnoff, 2011b, 2011a, 2012, 2013), presentational character (Bengson, 2015, p. 741), and phenomenology of givenness (Berghofer, 2020b, p. 162). The specific phenomenology is also credited to various philosophers in history: for example, Kriegel (2015a, p. 266) credits his view to Jean-Paul Sartre, Matthen (2005, p. 306) to David Hume (for a different interpretation of Hume's view, see Kriegel, 2015a; Siegel & Silins, 2015, p. 792), and Berghofer (2020b, p. 176) to Edmund Husserl. Husserl is, in turn, greatly influenced by Descartes. At the beginning of his *Cartesian Meditations: An Introduction to Phenomenology*, Husserl (1960) wrote: “[O]ne might almost call transcendental phenomenology a neo-Cartesianism” (p. 1).

me experiences] are very weak, i.e. they are not very assertive. Other seemings are so assertive that they make their contents feel utterly obvious. Right now it seems to me that $2 + 2 = 4$, that there is a desk in front of me, that I have a slight headache, and that dogmatism is true, but the latter seeming is far weaker than the former ones. (p. 530; see also Berghofer, 2020a, pp. 8-9; Berghofer, 2020b, p. 163; Huemer, 2005, p. 100)

Despite these similarities, it can be misleading, however, to equate phenomenal clarity with the assertiveness phenomenology. What exactly is the assertiveness phenomenology conceived by phenomenal dogmatists? In his proposal of phenomenal dogmatism, Pryor (2004) writes: “I called this the ‘phenomenal force’ of perceptual experience, thinking of it on analogy with the assertoric force of a public utterance.” (p. 372, footnote 23) In a similar vein, in her critique of phenomenal dogmatism, Teng (2018) writes that the assertiveness of perceptual experience is “analogous to the assertiveness of testimony.” (p. 641) Remarks like these from both the proponents and opponents of phenomenal dogmatism suggest that the assertiveness phenomenology of perceptual experience is analogous to the assertiveness phenomenology of testimony: in both cases p is *assertively* presented to S as true.

It seems to me, however, that a potential problem with this analogical suggestion is that the assertiveness phenomenology of testimony is not associated with phenomenal clarity: no matter how strongly a person asserts that p in their testimony, it is unlikely that it would hence become clear to the hearer that p . To see this point, consider Fermat's Last Theorem which is proved by Andrew Wiles. It is fair to say that it does not seem clear to most people that Fermat's Last Theorem is true (though it may have seemed clear to Fermat and a few other mathematicians as true). If by a strong assertion we mean that a proposition is asserted with a tone of confidence by a

trustworthy person, then let us suppose that we have the privilege to hear Wiles himself telling us that Fermat's Last Theorem is true. By hearing Wiles' assertion, we entertain the assertiveness phenomenology. The question is: would this strong assertion make it clear to us that Fermat's Last Theorem is true? It is doubtful, I think, that there would be such an effect. (By contrast, it may become clear to us that Fermat's Last Theorem is true with the help of Wiles' proof.)

Given that the assertiveness phenomenology of testimony is not associated with phenomenal clarity, if the suggestion were true that the assertiveness phenomenology of perceptual experience is the same as the assertiveness phenomenology of testimony (Pryor, 2004, p. 372, footnote 23; Teng, 2018, p. 641), then the assertiveness phenomenology of perceptual experience would not be associated with phenomenal clarity either. Is this the kind of phenomenology that concerns phenomenal dogmatists?

It is far from clear how phenomenal dogmatists would respond. Various responses might be developed depending on the details of their answers to the following questions: what exactly is the relationship between the assertiveness phenomenology of perceptual experience and the assertiveness phenomenology of testimony? In what sense are they analogous to each other and in what sense are they different? What is the relationship between the assertiveness phenomenology of perceptual experience and phenomenal clarity? Are they the same? Is there a difference between them? If so, what is it?

Here I shall not attempt to discuss possible answers on behalf of phenomenal dogmatists who are primarily concerned with perceptual justification.⁶⁷ By contrast, my present concern is how the assertiveness phenomenology can help us get a better grip on

⁶⁷ For example, Pryor (2004) writes: "My view is that our perceptual justification comes from that phenomenology. Having the phenomenology of seeming to ascertain P is what makes us have prima facie justification to believe P." (p. 357)

phenomenal clarity. For simplicity, I assume that when phenomenal dogmatists talk about the assertiveness phenomenology, they are talking about phenomenal clarity, but I suggest that phenomenal clarity is not the same as the assertiveness phenomenology of testimony. Though in both cases *p* is presented as true, they are distinct ways in which *p* is presented as true: whereas in a testimony *p* is presented as true in virtue of its assertiveness phenomenology, in a clear experience *p* is presented as true in virtue of its phenomenal clarity.

In this section, I have tried to elucidate what phenomenal clarity is by discussing what we are not talking about when we talk about phenomenal clarity. I hope the discussion helps narrow down the phenomenal character that concerns us when we move to discuss the central thesis that phenomenal clarity compels assent in the next section.

4. How does phenomenal clarity compel assent? An inquiry

According to the phenomenal reading of Descartes, it is evident that Descartes thought that phenomenal clarity compels assent (Patterson, 2008; Paul, 2020). But apart from being said by Descartes, why should we think that it is even possible that when a person has a clear experience in which it seems very clear to them that *p*, they would be compelled to believe that *p*? The possibility, I think, is based on two basic ideas. Very roughly, the argument is as follows:

- (1) We tend to believe that *p* if it seems clear to us that *p*.
- (2) The clearer it seems to us that *p*, the stronger our tendency to believe that *p* becomes.

Therefore,

- (3) It is possible that a person may have a clear experience in which it seems so clear to them that *p* that their tendency may reach the point where they are compelled to believe that *p*.

Regarding (1), it is obvious that we tend to believe that *p* if it seems clear to us that *p*; so obvious that it might strike some as if the claim about our tendency to form a belief and the claim about the phenomenal clarity of our experience are nothing but different descriptions of the same thing. This latter idea is, of course, not true. Among other reasons, it is not hard to conceive of cases in which a creature's tendency to believe that *p* has nothing to do with the way *p* is presented to them in their experience: for example, for this kind of creature, their clear experience that the sky is blue may have no impact on their tendency to believe that the sky is blue (and their belief that the sky is blue may be solely based on their independent scientific evidence).

(2) is a claim about the relationship between the degree of phenomenal clarity and the degree of the tendency to assent. It also seems quite intuitive. Besides being evident in the literature on the phenomenal reading of Descartes, it is also widely accepted in the literature on phenomenal dogmatism. For example, Huemer (2005) writes: "[W]e are more inclined to accept what more strongly seems to us to be true." (p. 100)

The combination of (1) and (2) suggests that (3) it is possible that a person may have a clear experience in which it seems so clear to them that *p* that their tendency may reach the point where they are compelled to believe that *p*. This possibility is, admittedly, not as intuitive as (1). One main reason may be that, while (1) is well illustrated by familiar cases, (3) is about extraordinary cases which are relatively rare in our daily lives. Another reason may be that, while we have quite a few explanations of (1), these explanations do not

immediately help us get a grip on how (3) is possible. For example, regarding (1), one may argue that it is because we are rational animals, and it is a rational requirement that rational animals tend to believe what seems clear to them. But, regarding (3), it does not seem to be a reasonable explanation that there is a rational requirement that we are compelled to believe what seems very clear to us.

The rest of this section aims to find a plausible explanation of how (3) is possible. Why should we care about this question? It does not seem to be a terribly exciting project if all it can do is to tell us that we do not merely tend to believe that the sky is blue when we clearly perceive that the sky is blue, but, under certain conditions, we are compelled to believe that the sky is blue.⁶⁸ It is a project that is worth exploring in detail because it may offer a potential explanation of baffling cases in which one is compelled to believe something in the face of counterevidence. The nature of this project is exploratory. I will examine a series of candidate accounts which, at first sight, may appear to be able to explain (1) and/or (2), and sometimes also (3) when the belief is true. But I will argue that most of them turn out to be unsuccessful in explaining the possibility that a person is compelled to believe what seems very clear to them in the face of counterevidence (Sections 4.1-4.2). In the end, I will develop an account according to which (3) is possible because, in borderline cases, the causal force of clear experience may outweigh the causal force of evidence even when the justificatory force of clear experience is undermined by the justificatory force of evidence. (Section 4.3).

4.1. Austin's pig

One influential way to appreciate the distinctiveness of believing what seems clear to us is to compare it with believing what

⁶⁸ Unlike Moore's project (1939/2013), this project does not help us deal with scepticism.

evidence supports. The contrast between them is manifest in the widely-quoted passage by Austin (1962):

The situation in which I would properly be said to have evidence for the statement that some animal is a pig is that, for example, in which the beast itself is not actually on view, but I can see plenty of pig-like marks on the ground outside its retreat. If I find a few buckets of pig-food, that's a bit more evidence, and the noises and the smell may provide better evidence still. But if the animal then emerges and stands there plainly in view, there is no longer any question of collecting evidence; its coming into view doesn't provide me with more evidence that it's a pig, I can now just see that it is, the question is settled. (p. 115)

In the first case, Austin has an evidence-based belief: his belief that there is a pig is based on his evidence such as the pig-like marks, the pig food, the noises, and the smell. In the second case, Austin has a perceptual belief: his belief is not based on his aforementioned evidence but is acquired when he sees that the pig “stands plainly in view” (p. 115). Obviously, clearly perceiving that there is a pig and having evidence related to a pig are quite different: the former can, while the latter cannot, *settle* the question of whether there is a pig. But how can clearly perceiving a pig settle the question? There might be several different answers depending on our reading of Austin's passage. In the following, I will discuss three different readings about what it is that settles the question, and I will argue that all three readings are not able to account for how it is possible that a person is compelled to believe what seems very clear to them in the face of counterevidence.

4.1.1. Being in primitive contact with reality

One reading of Austin's passage is that by talking about how clearly perceiving that p settles the question of whether p , Austin is mainly talking about a fundamental way we acquire knowledge: that is, by being in primitive contact with reality (see e.g. Ayers & Antognazza, 2019; Antognazza, 2021). Following this reading, to say that a pig stands plainly in view is just a way to say that the perceiver is in primitive contact with reality. This may well be the intended reading by Austin, especially when we consider the broader tradition of Oxford Realism (Longworth, 2020).

This reading does not need to say anything about the phenomenal character of being in primitive contact with reality. As far as this reading is concerned, it may not matter a lot whether the experience in which we are in primitive contact with reality has phenomenal clarity or not. If the latter point is true, then this reading does not suffice to help address our concern about the relationship between phenomenal clarity and belief.

4.1.2. The phenomenal character of being in primitive contact with reality?

By contrast, one may read Austin's passage as talking about the phenomenal character of the experience of being in primitive contact with reality. According to this reading, to say that the pig stands plainly in view is just another way to say, in our terminology, that the perceiver has a clear experience in which it seems clear to the perceiver that there is a pig.

Furthermore, a drastic version of this reading may argue that the phenomenal clarity of the experience is an *infallible* maker of being in primitive contact with reality (see e.g. Carriero, 2013, p. 195). This is also a plausible interpretation of Descartes' notion of clarity when

he proposed that whatever he clearly and distinctly perceives is true. If it were true that phenomenal clarity is an infallible marker of being in primitive contact with reality in the sense that whatever seems clear to the subject is in fact true, then no wonder the question of whether it is true is settled.

This conception of phenomenal clarity, however, seems implausible. This is because phenomenal clarity is a form of subjective experience, and it is implausible that a certain form of subjective experience can infallibly settle the question of whether the experienced is true. Neither Descartes (see e.g. Patterson, 2008) nor anyone in the history of philosophy and beyond has ever succeeded in arguing that there is such a distinctive phenomenal character. To be clear, here I am not suggesting that the question of whether *p* can never be settled, but only that no matter what it is that settles the question, it is unlikely to be the phenomenal character of the subject's experience.

4.1.3. Knowing that *p* versus merely thinking that one knows that *p*?

There may be a third reading of Austin's passage, which is weaker in comparison with the second reading. It is concerned with the question of why the perceiver *thinks* that the question of whether there is a pig is settled, rather than why the question of whether there is a pig is settled as a matter of fact. According to this third reading, the settlement is the result of a commonsensical psychological process. What exactly is the commonsensical psychological process? In the literature, there is a line of thoughts according to which perceiving that *p* is a typical way of knowing that *p* and knowing that *p* entitles the perceiver to believe that *p* (Roessler, 2019; see also Cassam, 2007; Millar, 2011; Williamson, 2000). The third reading takes inspiration from this view; but instead of focusing on factive

perception, the third reading suggests that, although a clear experience in which it seems to the subject that *p* is no guarantee that it is a factive perception, it is common practice for people to take a clear experience as a factive perception: that is to say, no matter whether the subject *in fact* perceives that *p*, in terms of their clear experience they may *think* that they in fact perceive that *p*, and *think* that they therefore know that *p*. Given that the perceivers *think* that they know that *p*, the question of whether *p* is settled for them.

This reading can explain why (1) people tend to believe what seems clear to them: it is because they think that their clear experience is a factive perception. If it is also true that the clearer it seems to the subject that *p*, the stronger their tendency to take their clear experience as a factive perception becomes, then this reading can also explain why (2) the clearer it seems to the subject that *p*, the stronger their tendency to believe that *p* becomes.

This reading, however, does not seem to be able to explain how it is possible that a person is compelled to believe what seems very clear to them in the face of the evidence that the content of their clear experience is not true. This is because no matter how much their tendency to believe the content of their clear experience is, it is based on the assumption that their clear experience is a factive perception. Once the assumption is proved to be wrong in the face of counterevidence, their tendency to believe that *p* should collapse with it.

To recap, Austin's famous passage draws our attention to the distinctiveness of believing what we clearly perceive, as opposed to believing what evidence supports: clearly perceiving that *p* can settle the question of whether *p*. While Austin may well be correct about the settlement force of clear perception in the pig case, our discussion of the three readings suggests that the reason why clear perception has the settlement force in the pig case may not suffice to account for

cases in which the subject is compelled to believe what seems very clear to them in the face of counterevidence. To explain how the latter cases are possible, we need to look for factors that are not immediately salient in the pig case.⁶⁹

4.2. Justification

We have previously discussed the similarities between phenomenal clarity and the assertiveness phenomenology concerning phenomenal dogmatists (Section 3.6), and I have suggested an interpretation according to which, when phenomenal dogmatists talk about the assertiveness phenomenology, they are talking about phenomenal clarity. Like our present concern about the relationship between phenomenal clarity and belief, phenomenal dogmatists are also concerned with the relationship between the assertiveness phenomenology and belief. This makes phenomenal dogmatists' view on the relationship between the assertiveness phenomenology and belief a natural candidate explanation of why phenomenal clarity compels assent.

On the face of it, phenomenal dogmatists may have an explanation of why (1) we tend to believe that p if it seems clear to us that p : it is because, according to phenomenal dogmatists, such an experience offers "immediate (*prima facie*) justification for believing p " (Pryor, 2000, p. 532).

Phenomenal dogmatists may also have an explanation of why (2) the clearer it seems to us that p , the stronger our tendency to believe that p becomes: it is because, according to phenomenal dogmatists, the degree of justification offered by an experience covaries with the degree of its assertiveness phenomenology.

⁶⁹ There is a fourth reading of Austin's passage, according to which the settlement force of clear perception is understood as a form of decisive causal force (Campbell, 2014, pp. 80-81; for critique, see Cassam, 2014, pp. 133-134). I will discuss the causal point in Section 4.3.

Berghofer (2020b) writes: “Experiences do not either have or not have a justification conferring phenomenology, they can have it in a more or less pronounced way. The more pronounced, the more justification they provide.” (p. 163; see also Berghofer, 2020a, pp. 8-9) If it is true that the clearer it seems to that p, the more justification we have for believing that p, then it makes sense that with more justification for believing that p, we have a stronger tendency to believe that p.

Given that the combination of (1) and (2) suggests that (3), we might expect that, by appealing to the justification offered by experience, phenomenal dogmatists can explain (3) as well. It is, however, doubtful that this could work.

The main reason is that although it is arguable that our justification for believing that p may increase indefinitely with more and more evidence that p (see e.g. Nelson, 2002), it is unlikely that the justification offered by the experience can increase indefinitely. Crucially, it is unlikely that the justification offered by the experience can increase to a point where the subject cannot but believe that p even in the face of the evidence that not-p. This is because when a person has the evidence that not-p, it will undermine or outweigh the justification offered by the experience. This point is evident in the following paragraph by Pryor (2000):

This *prima facie* justification can be undermined or threatened if you gain positive empirical evidence that you really are in a skeptical scenario. (For instance, if a ticker tape appears at the bottom of your visual field with the words “You are a brain in a vat...”) If you acquire evidence of that sort, then you’d have to find some non-question-begging way of ruling the skeptical hypothesis out, before you’d be all things considered justified in believing that things are as your experiences present them. (pp. 537-538)

Here, Pryor does not consider the possible increase of phenomenal clarity. But it seems a plausible view that, in the light of the evidence that not-*p*, we would no longer be justified in believing that *p* no matter how clear it seems to us that *p*.

It is worth noting that this is not to say that there is absolutely no covariation between the justification offered by the experience and the phenomenal clarity of the experience. I do think that there is such a covariation but only within a very limited range. Within this limited range, the clearer it seems to the subject that *p*, the more justified they are to believe that *p*. But there is a threshold on how much justification a clear experience can offer. And the threshold is likely to be very low. After reaching the threshold, the justification offered by the experience would not covary with the phenomenal clarity of the experience.

4.3. Causation

So far, we have discussed four attempts to explain how (3) is possible. All four attempts are developed based on well-established philosophical theories. But none of them turned out to be successful. One might hence wonder whether this suggests that such a possibility does not exist. I think this suggestion is too quick, for several reasons: first, the lack of a satisfactory theory of how it is possible itself does not entail that it is impossible; second, (3) is based on two very intuitive premises, and we do not have an immediate reason to doubt either premise; third, as I will argue, there is a causal conception of the relationship between phenomenal clarity and belief which may explain how (3) is possible.

According to the causal conception, we can explain the transition from a clear experience in which it seems clear to the subject that *p* to the belief that *p* in terms of a form of psychological

causation between two mental states.⁷⁰ This causal conception has not received adequate attention in the scientific literature or in the philosophical literature in general. But it can be discerned in the philosophical literature on phenomenal dogmatism.

Although the primary concern of phenomenal dogmatists is justification, some of them appear to have mentioned the causal relationship between phenomenal clarity and belief. For example, Huemer (2005) writes: “[W]e are more *inclined* to accept what more strongly seems to us to be true.” (p. 100, emphasis added) Koksvik (2011) writes: “[Perceptual experience] has phenomenology of *pushiness* when its *pushing* its subject to accept its content is itself an aspect of its character.” (p. 260, emphasis added) Here, the terms *incline* and *push* seem to be about the causal force of perceptual experience.

Opponents of phenomenal dogmatism may accept the causal conception as well. For example, in her critique of phenomenal dogmatism, Teng (2018) argues that “it is possible for imaginings to

⁷⁰ In the following quotation, Descartes (1984) appeared to talk about two sorts of perception: clearly perceiving something and perceiving a manifest contradiction:

[W]hen I turn to the things themselves *which I think I perceive very clearly*, I am so convinced by them that I spontaneously declare: let whoever can do so deceive me, he will never bring it about ...that two and three added together are more or less than five, or anything of this kind *in which I see a manifest contradiction*. (p. 25, emphasis added)

This invites the question of how we should understand the relationship between clearly perceiving that p and the absence of perceiving a manifest contradiction when one perceives that p: are they the same? Or are they necessarily associated with each other in the sense that whenever you clearly perceive something, you would not perceive a manifest contradiction, and vice versa? Or is the absence of perceiving a manifest contradiction just an additional but dissociable reason for one to believe what they clearly perceive?

It is not clear to me what Descartes' answer was. Nevertheless, all I am proposing is that phenomenal clarity may be a causal factor. I surely do not deny that there are other factors in belief formation: needless to say, the evidence that p is another important factor for us to believe that p. For simplicity, I shall set aside complications concerning the relationship between clearly perceiving that p and the absence of perceiving a manifest contradiction when one perceives that p.

have the distinctive phenomenal character dogmatists have in mind.” (p. 637; see also Harrison, 2019) By the distinctive phenomenal character, she means the phenomenal character that offers justification. She does not explicitly talk about the causal force of the distinctive phenomenal character. Her main point is that the same phenomenal character is present in both veridical sensory perception and imaginings. However, if it is correct that the distinctive phenomenal character of sensory perception has causal force, then Teng’s argument also supports the view that imaginings can have the same distinctive phenomenal character with causal force. In a similar vein, Chasid and Weksler (2020) argue that opponents of phenomenal dogmatism “should be willing to accept the minimal characterization of perceptual assertoricity as the property of *inclining* the perceiver to believe the content of her experience” (p. 733, emphasis added)

The general point is that, though there is disagreement about whether a clear experience, in which it seems clear to the subject that *p*, offers immediate justification for believing that *p*, both the proponents and opponents of phenomenal dogmatism may accept the causal conception of the relationship between phenomenal clarity and belief.⁷¹

What kind of causation are we talking about here? I think the kind of causation between phenomenal clarity and belief is a *primitive* relationship between two mental states. It is primitive in the sense that whether a person’s clear experience that *p* would cause their belief that *p* does not depend on (the person’s consideration of) whether

⁷¹ For Descartes, clear perception does not only have causal force, it is also a guide to, or even a guarantee of, truth: when one’s perception is *completely* clear, “it’s infallible.” (Paul, 2020, p. 12) It is, however, important to note that the causal conception is not committed to the latter claim about the truth or infallibility of clear perception. That is to say, even though there is a general agreement that Descartes did not succeed in arguing for the infallibility of completely clear perception (Patterson, 2008), the causal conception of the relationship between phenomenal clarity and belief remains plausible.

clear experience offers justification for believing that p. Call it the primitive view on the causal force of clear experience.

This view is arguably held by many philosophers who may disagree about other aspects of the nature of experience. For example, in his argument for naïve realism, Campbell (2014) discusses both the causal force and the justificatory force of perceptual experiences. He writes:

I will put the causal point by saying that perceptual experiences are decisive in the formation of beliefs about one's surroundings. I will put the normative point by saying that perceptual experiences are authoritative in the formation of beliefs about one's surroundings. (p. 81)

Importantly, Campbell's conception of causation is based on the notion of intervention, rather than on the notion of justification (see e.g. Campbell, 2007).⁷² In this sense, Campbell's work can be read as being in line with the primitive view on the causal force of experience.

Furthermore, Campbell (2014) also appears to have briefly considered the possible dissociation between the causal force and the justificatory force of perceptual experiences in delusions:

[I]n the case of certain delusional patients, ... prior beliefs overwhelm the input from perception. But it does not happen in ordinary humans, just because of the decisive [causal] role played by perceptual experience. (p. 84)

In my understanding, Campbell's main point is that, though, in ordinary humans, perceptual experience plays the decisive causal

⁷² It is debatable whether the causal force of perceptual experience is always decisive in everyday cases (see e.g. Cassam, 2014, pp. 133-134). Here I shall not take a stand on this debate. I assume it is relatively uncontroversial that perceptual experience has primitive causal force, and I suggest that it can be decisive in cases like akratic beliefs and delusions.

role in belief formation, it can lose its decisive causal force in some cases of delusions. Would perceptual experience also lose its justificatory force in such cases of delusions? It is not entirely clear what Campbell's view is. But it seems to be a plausible view that perceptual experience may retain its justificatory force even in such cases of delusions: that is to say, the patients are still not justified to hold the delusional beliefs, even though the causal force of perceptual experience is overwhelmed by prior beliefs. If this analysis is along the right lines, then here Campbell is talking about a dissociation between the causal force and justificatory force of perceptual experience in such cases of delusions.⁷³

The dissociation between the causal force and justificatory force of perceptual experience is also evident in the following passage by Davidson (1986) who is not a naïve realist:

The relation between a sensation and a belief cannot be logical, since sensations are not beliefs or other propositional attitudes. What then is the relation? The answer is, I think, obvious: the relation is causal. Sensations cause some beliefs and in this sense are the basis or ground of those beliefs. But a causal explanation of a belief does not show how or why the belief is justified. (p. 310)

Though it is debatable whether Davidson is correct that experience cannot justify belief (see e.g. BonJour, 1985; Brewer, 1999; McDowell, 1994; Pryor, 2005), I think Davidson's point is in line with the primitive view that the causal relationship between experience and belief is primitive, and it is not dependent on the justificatory

⁷³ Campbell is concerned with the causal force of veridical perception that constitutes a great part of the evidence that the patients' delusional beliefs are not true. He does not consider the causal force of the patients' clear experiences in which it seems very clear to the patients that their delusional beliefs are true. I shall focus on the latter kind of experience.

relationship between them. In what follows, I shall assume that the primitive view is true.

The mere fact that the causal relationship between clear experience and belief is primitive still does not suffice to explain how it is possible that a person is compelled to believe what seems very clear to them in the face of counterevidence. To explain the possibility, we need a more detailed account of the interaction between the causal force and justificatory force of the clear experience that p , on the one hand, and the causal force and justificatory force of the evidence that $\text{not-}p$, on the other hand.

4.3.1. The causal force and justificatory force of clear experience

At the end of Section 4.2, I have suggested that the justification offered by a clear experience cannot increase indefinitely: there is a threshold on how much justification a clear experience can offer, and the threshold is likely to be very low. Given that justification and causation are often assumed to go hand in hand in many areas of practical and theoretical reasoning, one might suggest that there is a threshold on how much causal force a clear experience can have, and the threshold on the causal force of a clear experience should be in line with the threshold on its justificatory force, both of which are very low.

I think this suggestion is wrong. The main motivation behind the assumption that justification and causation go hand in hand is the fact that, when we consider the relationship between evidence and belief, the more evidence we have for believing that p , the more justification we have, and the more we are causally inclined to believe that p . I, of course, accept this fact. But it is important to note that, *unlike the primitive causal force of clear experience, which is independent of its justificatory force, the causal force of evidence is*

not primitive: the causal force of evidence is not independent of the justificatory force of evidence; rather, how much causal force the evidence has is determined by how much justification the evidence offers. In other words, the reason why, in the light of the evidence, we are, to a certain degree, causally inclined to believe that *p* is precisely that the evidence justifies us to believe that *p* to such a degree. Since the primitive causal force of clear experience does not depend on its justificatory force, we do not have the main motivation for accepting that, when the justificatory force of a clear experience reaches its threshold, the causal force of the clear experience would stop increasing. At the very least, it seems possible that, as the phenomenal clarity of an experience increases, while the justification may soon reach the threshold and the clear experience stops offering additional justification, its primitive causal force may continue to increase.

This paints a picture in which justification and causation come apart. Is it a plausible view that justification and causation can come apart? I think it is. The view is not entirely new. In the literature on practical reasoning, Davidson (1963) has already made the point that a person may have several reasons for an action but it is possible that only one of the reasons is the cause of the action (see also Campbell, 2020, pp. 9-10, 60). More pertinent to our concern may be Davidson's (1969) account of *akrasia*, according to which the causally strongest reason differs from the reason deemed by the agent to offer the strongest justification. What I have been trying to argue is that justification and causation may come apart in the theoretical domain as well.

4.3.2. The causal force and justificatory force of evidence

As our evidence for believing that *p* accumulates, our justification for believing that *p* may increase indefinitely (see e.g.

Nelson, 2002). Given that the causal force of evidence is not primitive and depends on the justificatory force of evidence,⁷⁴ one may wonder if the causal force of evidence may increase indefinitely as the evidence accumulates.⁷⁵

I think it is a plausible view that the causal force of evidence will *not* increase indefinitely and can only increase to a certain point, a candidate for which is the point where there is sufficient evidence for believing that *p*. Considering the following case:

PANDEMIC: In early 2020, many people, including the WHO, were not very much inclined to believe that there is a pandemic. When there was more and more evidence, people had more and more justification for believing that there is a pandemic and were more and more inclined to believe so. At a certain point, there was sufficient evidence that there is a pandemic. After that point, the evidence still accumulates every day, and our justification for the belief also increases every day. The question is whether we continue to be more and more causally inclined to believe that there is a pandemic.

It seems to me that our inclination to believe in the existence of the pandemic reached its peak once we had sufficient evidence that there is a pandemic, even though our justification continued to increase.

⁷⁴ By justification, I mean propositional justification rather than doxastic justification. Propositional justification concerns whether a person has good evidence to believe a proposition, whereas doxastic justification concerns whether a person's belief is based on good evidence. For discussion, see e.g. Turri (2010), and Silva Jr and Oliveira (forthcoming).

⁷⁵ The claim that evidence has causal force is compatible with both doxastic involuntarism which denies that we can believe at will (e.g. Williams, 1973; Alston, 1988), and doxastic voluntarism which argues that at least in *some* cases we have direct voluntary control over our beliefs (e.g. Ginet, 2001; Weatherson, 2008; Turri et al., 2018)

Similarly, consider whether we are more inclined to believe that a mathematical equation is true when we have three hundred ways to prove that it is true, as opposed to when we have only one way to prove that it is true. With two hundred and ninety-nine extra proofs, we arguably have more justification for believing that the mathematical equation is true. But it seems to me that we would not be *more* causally inclined to believe that the mathematical equation is true.

4.3.3. How is it possible that a person cannot but believe something in the face of counterevidence?

If the above analysis is on the right track, then there is an asymmetry between the way the causal force and justificatory force of clear experience come apart and the way the causal force and justificatory force of evidence come apart. Regarding clear experience, within a limited range, as the clarity increases, the justificatory force increases hand in hand with the causal force; but after a certain point, as the clarity keeps increasing, the justificatory force may reach its threshold while the causal force may continue to increase. Regarding evidence, within a limited range, as the evidence accumulates, the causal force increases hand in hand with the justificatory force; but after a certain point where the evidence is sufficient, as the evidence keeps accumulating, the causal force may reach its peak while the justificatory force may continue to increase.

This asymmetry makes it possible that there are cases in which a person has a very clear experience with enormous causal force but limited justificatory force, on the one hand, and lots of counterevidence with enormous justificatory force but only limited causal force. In such cases, the justificatory force of the counterevidence outweighs the justificatory force of the clear experience, and the person hence is not justified to believe what

seems very clear to them; but the causal force of the clear experience may outweigh the causal force of the counterevidence, and the person is hence causally compelled to believe what seems very clear to them.^{76, 77}

5. Proposing a dual-force framework of beliefs

There are two fundamental questions about the aetiology of a belief: one is about whether the belief is justified; the other is about what causes the belief. There are also two important factors related to justification and causation: one is clear experience; the other is evidence. In the literature, it is sometimes said that clear experience has phenomenal force and evidence has evidential force. Depending on the context, the term *phenomenal force* can mean either clear experience's justificatory force or its causal force; and the term *evidential force* can mean either evidence's justificatory force or its causal force. The phenomenal force and the evidential force constitute what we can call *the dual-force framework* of beliefs, according to which many beliefs are the results of the interaction

⁷⁶ I take it that this possibility only happens in extraordinary cases, and it is not at odds with ordinary cases in which the perceiver is able to decide whether to “take the experience at face value.” (McDowell, 1994, p. 26). In my understanding, cases like the Müller-Lyer illusion do not count as extraordinary cases, and the phenomenal clarity does not reach the point to compel assent in the light of counterevidence. This is why the perceiver can “refrain from judging that that is how things are.” (McDowell, 1994, p. 11, footnote 9)

⁷⁷ One might wonder whether the counterevidence can simply undercut the causal force of phenomenal clarity. I think that there is a sense that the presence of the counterevidence may not simply outweigh but undercut or eradicate the justification offered by the clear experience. But I see no obvious reason how the causal force can be undercut, especially considering the causal force of experience is primitive whereas the causal force of evidence is not. I also assume that certain qualities of clear experience may persist in the light of counterevidence. Similar views are also held by many phenomenal dogmatists when they talk about the assertiveness phenomenology. For example, Brogaard and Gatzia (2017) write: “[T]he mark of justifying experiences is ... the feeling that the experience is so solid that it would not disappear even if we were to discover that it is non-veridical.” (p. 545; see also Brogaard, 2013, p. 275)

between the phenomenal force of clear experience and the evidential force of evidence.

In optimal cases, S may have abundant evidence that p, and S may also have a clear experience in which it seems clear to S that p. S is hence both evidentially and phenomenally forced to believe that p. More specifically, S's evidence justifies the belief that p and causally inclines S to believe that p; S's clear experience provides additional justification for believing that p and also causally inclines S to believe that p.

In less optimal cases, S may have a clear experience in which it seems clear to S that p, but at the same time S may also have some evidence that not-p. Whether S would be justified to believe that p depends on whether the justification offered by the clear experience can outweigh the justification offered by the evidence that not-p; and whether S would believe that p depends on whether the causal force of the clear experience can outweigh the causal force of the evidence that not-p.

In the Müller-Lyer illusion in which it seems, to a certain degree, clear to S that the two lines have different lengths, S is not justified to believe that the two lines have different lengths when S has sufficient evidence that the two lines have the same length, because the justification of S's evidence outweighs or undercuts the immediate justification offered by S's clear experience. S would believe that the two lines have the same length because the causal force of the evidence outweighs the causal force of the clear experience.

In more extraordinary cases, S may have a very clear experience in which it seems very clear to S that p such that S is compelled to believe that p even in the face of the evidence that not-p. As I explained in the last section, in such cases, S is not justified to believe that p, because the justification of the evidence that not-p outweighs or undercuts the immediate justification offered by the clear experience that p; but S is causally compelled to believe that p,

because the causal force of the clear experience that p outweighs the causal force of the evidence that not-p.

The dual-force framework paints a complicated picture of belief formation, in comparison with what can be called a singular evidential-force framework of belief formation, according to which the only factor in the aetiology of beliefs is the justificatory force of evidence. Admittedly, the singular evidential-force framework can offer simplistic explanations of the optimal cases and cases like the Müller-Lyer illusion: in both cases S believes that p simply because it is what S's evidence justifies. Should this be a reason to favour the singular evidential-force framework over the dual-force framework? I think here simplification may not be a virtue because it clouds what is actually going on in optimal cases and less optimal cases like the Müller-Lyer illusion. In this sense, this chapter has been trying to reconstruct the complicated picture of belief formation. This way of philosophising as reconstructing rather than simplifying is also pronounced in Hyman's work on agency. In the preface of his book *Action, Knowledge, and Will*, Hyman (2015) writes: "One aim of the book is to criticize this relentlessly simplifying philosophy and reconstruct the edifice, separating the storeys that were sandwiched together in the collapse." (p. ix)

More importantly, while the singular evidential-force framework struggles in explaining how the extraordinary cases are possible, the dual-force framework can offer a plausible explanation. The explanatory power of the dual-force framework is another reason for us to accept that this complicated picture is close to what is actually going on in the aetiology of beliefs. The explanatory power of the dual-force framework in explaining how it is possible for clear experience to compel assent in the face of counterevidence gives us a motivation for examining whether it can help us explain some more concrete cases in which one believes something in the face of counterevidence.

Apart from delusions, akratic beliefs are also notorious for their flouting of evidence, albeit to a relatively lesser degree.⁷⁸ Before we dive into the details about how the dual-force framework can help explain delusions, it may be worth testing the explanatory power of the dual-force framework by focusing on examining how it can help us explain akratic beliefs. If it turns out that the dual-force framework can offer a better explanation of akratic beliefs, then it will give us additional reason to accept the dual-force framework, and additional motivation for examining how it can help us explain delusions, which are arguably the most baffling beliefs.

6. The dual-force account of akratic beliefs

6.1. A minimal conception of akratic beliefs

To explain akratic beliefs, it is essential to be clear about what counts as an akratic belief. Akratic belief is not a familiar term in the ordinary language. It is devised by philosophers to refer to a unique kind of baffling doxastic phenomenon. What exactly counts as an akratic belief? The answers provided by philosophers differ in detail. But there is a somewhat minimal conception of akratic belief which seems to be shared in the literature. According to this minimal conception, S akratically believes that p, when the following conditions obtain:

- (A) The belief condition: S believes that p.
- (B) The evidence condition: S believes that her evidence shows that not-p.

⁷⁸ In this thesis, I shall set aside the thorny issue about the necessary and sufficient conditions for distinguishing an akratic belief from a delusional belief. Instead, my strategy is to focus on the typical examples of akratic beliefs and delusional beliefs in the literature.

(C) The concurrent consciousness condition: S is consciously aware of (A) and (B) at the same time.

It is minimal in comparison with several popular conceptions of akratic belief. In some philosophers' writings, the evidence condition (B) is replaced by what we can call the rationality or normativity condition: that is, S believes that it is irrational to believe that p, or normatively required not to believe that p (e.g. Heil, 1984; Hookway, 2001; Kearl, 2019; Smithies, 2019). The evidence condition (B) is less demanding than the rationality or normativity condition, because even though the consideration of evidence is an important part of the consideration of whether a belief is rational or normatively required, the latter may include considerations beyond evidence. For example, whether a belief is rational may also depend on its relationship with the subject's action (see e.g. Hookway, 2001).

Another popular conception of akratic belief adds the following requirement (Owens, 2002; see also Tenenbaum, 1999):

(D) S freely and deliberately forms the belief that p.

It is not easy to get a full picture of what exactly (D) asks for. In her critique of (D), Tanney (2017) writes: "[I]t is not even clear what would count as 'freely and deliberately' or 'freely and intentionally' believing that something is so." (p. 353) It looks to me, though, at least one important aspect of Owen's (2002) conception is that:

For such a[n] [akratic] belief to be formed freely and deliberately, the agent must be in a position to judge that a certain bit of evidence provides some reason for the belief, whilst also judging that this evidence is decisively outweighed by other evidence. (p. 390)

This quotation seems to suggest that (D) is virtually adding another condition to the minimal conception:

(D*) S believes that a small portion of her evidence supports the view that p.

The minimal conception, however, does not require that S has a small portion of evidence supporting her belief. Therefore, it does not require (D*) or (D).

The third popular conception of akratic belief is concerned with the relationship between akratic belief and akratic action (e.g. Adler, 2002; see also Setiya, 2013). It adds the following requirement:

(E) There is a structural similarity between akratic belief and akratic action.

While the relationship between akratic action and akratic belief is an important issue, it is beyond my present concern and the minimal conception does not require (E).

In short, the minimal conception of akratic belief does not address issues related to rationality, normativity, freely and deliberately believing, or akratic action. I imagine that one might wonder whether it is too minimal, and argue that some of the latter elements are constitutive of akratic belief. In response, I do not claim that I am giving a complete definition of akratic belief. Rather, my claim is that if we want to explain how an akratic belief is possible, we first need to be clear about what counts as an akratic belief and the minimal conception provides a starting point.

I also imagine that one might wonder whether the minimal conception is too demanding.⁷⁹ The necessity of the belief condition (A) and the evidence condition (B) is obvious. Let us briefly consider why the concurrent consciousness condition (C) is also necessary to

⁷⁹ For example, Smithies is explicit that he talks about a form of epistemic akrasia which is not “open eyed” (Smithies, 2019, pp. 287-288). Roughly, to say that an akratic belief is “open eyed” is to say that it meets the concurrent consciousness condition (C). Therefore, Smithies is not talking about a kind of belief which meets the minimal conception of akratic belief.

capture the phenomenon of akratic belief. (C) requires that S is consciously aware of (A) and (B) at the same time. Without (C), the minimal conception only requires that (A) S believes that p and (B) S believes that their evidence shows that not-p. There are a wide range of phenomena that meet (A) and (B). First, consider a typical belief-updating process: at time t1 I believed that p; then at time t2 I acquired lots of evidence that not-p and consequently I stopped believing that p. In this belief-updating process both (A) and (B) are met, despite that they are met at different times. But this belief-updating process is not what philosophers have in mind when they talk about akratic belief.

Second, for many philosophers (C) is what distinguishes epistemic akrasia from self-deception. Hookway (2001), for example, argues that self-deception does not meet (C). He writes: “[W]hen the jealous man self-deceptively believes that his partner is being unfaithful to him, he *cannot be, at the same time, aware* that he is utterly unreasonable so to believe.” (p. 181, emphasis added)

Third, (C) helps distinguish akratic beliefs from a certain conception of inconsistent beliefs. According to Davidson (1985/2004), inconsistent beliefs can co-exist in different fragments of the subject’s mind but cannot be present in the subject’s consciousness at the same time. Suppose in one fragment, f1, of S’s mind, S holds a belief that p and the evidence, e1, supporting that p, while in another fragment, f2, of S’s mind, S holds a belief that not-p and the evidence, e2, supporting that not-p, S would count as holding inconsistent beliefs. Now consider S’s belief that p, even if (A) is arguably met for S believes that p in f1 and (B) is arguably met for S believes that e2 shows that not-p in f2, S’s belief that p is not an akratic belief. Why? The main reason is that (C) is not met, since S’s belief that p and S’s evidence, e2, supporting that not-p are in different fragments of S’s mind, rather than being present to S’s consciousness at the same time. To be clear, I do not claim that it is

impossible for S to hold the belief that p and the evidence, e2, supporting that not-p in the same fragment of her mind. In such a case, S's belief that p might be taken as similar to an akratic belief. But such a case is not what Davidson has in mind when he talks about inconsistent beliefs.

With a minimal conception of akratic belief in hand, we have something, as a start, to examine various accounts of how an akratic belief is possible. The core proposal of the dual-force account of akratic belief is that S's akratic belief that p is the result of the interaction between S's clear experience in which it seems very clear to S that p, on the one hand, and S's evidence that not-p, on the other hand: S is not justified to believe that p because the justificatory force of S's evidence that not-p outweighs or undercuts the justificatory force of S's clear experience that p; but S is causally compelled to believe that p because the causal force of S's clear experience that p outweighs the causal force of S's evidence that not-p.

In the following, I will focus on developing the dual-force account of akratic beliefs, with the help of paradigmatic examples of akratic beliefs, and by comparing it with, and taking in important insights from, several current accounts of akratic beliefs. I will argue that the dual-force account can help make some progress in explaining akratic beliefs.

6.2. Current accounts of akratic beliefs versus the dual-force account

6.2.1. Reason and attention

Consider the following case in which Scanlon (1998, p. 35) talks about his own akratic belief:

FRIEND: (A) Scanlon believes that Jones is a loyal friend, despite that (B) he also believes that his evidence shows that Jones is an artful deceiver. And (C) he is consciously aware of both at the same time.

The belief that Jones is a loyal friend is an akratic belief. But how is it possible? Scanlon's own explanation seems to be that his seeming experience in which Jones appears to be a loyal friend is his reason for akratically believing that Jones is a loyal friend. Scanlon writes:

[I]n the case of the false friend, mentioned above, there is something that I take to be a reason for believing in his genuineness, namely his appearance of genuineness. Given all that I know about him, of course, I know that this is not a good reason in this case, but it can serve as my reason nonetheless. (p. 36)

The problem with this explanation is, however, that it is not clear why one takes what they know is not a good reason as their reason, especially considering the fact that we normally do not take what we know is not a good reason as our reason. For some philosophers, this is not even possible. For example, Raz (2009) writes: "[T]here is no possibility of preferring to follow what one takes to be the lesser reason rather than the better one." (p. 42)

Mele (1986, p. 216) suggests a different explanation of epistemic akrasia, according to which for the akratic subject the evidence that *p* is more *salient* than the evidence that not-*p*, in the sense that it grabs the subject's *attention* when they consider what to believe. Mele writes:

[I]tems which he takes to provide only weak evidence for *P* may, due to his conative condition, have a great deal of salience at the time at which he forms or

acquires a belief about the matter. Because of the attractiveness of what they suggest, the bulk of his attention may be drawn to these items, and his apprehension of competing items may be quite pale by comparison. (p. 216)

Following this explanation, one might think that the subject akratically believes that Jones is a loyal friend, because the subject's attention is grabbed by the salience of the seeming experience in which it seems that Jones is a loyal friend.

The problem with this attention-grabbing explanation is, however, that it seems to be directly at odds with (C) the concurrent consciousness condition: the subject is consciously aware of (A) and (B) at the same time, in which case the subject arguably should pay equal and fair attention to all evidence. Indeed, Mele accepts that it would be "difficult for [the akratic subject]... to retain the belief if we manage to shift the focus of his attention to his reasons for not believing that P." (p. 218) Therefore, the attention-grabbing explanation is not able to explain the minimal conception of akratic belief.

I think part of Scanlon's insight is correct that the seeming experience in which Jones appears to be a loyal friend is an important factor in the aetiology of the akratic belief; I also think part of Mele's insight is correct that the way the seeming experience affects the subject is an important factor. The dual-force framework offers an alternative suggestion to develop these insights. According to the dual-force framework, the seeming experience does not act as the subject's reason for the akratic belief (rather, the subject is consciously aware that the seeming experience does not count as a reason to justify the akratic belief in the light of the counterevidence); nor does the seeming experience affect the subject by shifting the subject's attention away from the counterevidence. Instead, the role of the seeming experience is to exert a primitive causal force that

compels assent. In the case of FRIEND, Scanlon akratically believes that Jones is a loyal friend, because Jones is such an artful deceiver that he succeeds in inducing Scanlon's clear experience whose primitive causal force outweighs the causal force of Scanlon's evidence, and hence causally compels his assent.

6.2.2. Practical consideration

Consider the following case of epistemic akrasia, adapted from Hookway (2001):

MOTHER: (A) A mother believes that her son is innocent, despite that (B) she also believes that her evidence shows that her son is not innocent. And (C) she is consciously aware of both at the same time.

The mother's belief that her son is innocent is an akratic belief. But how is it possible? Hookway (2001) argues that there are two conflicting evaluations in the case of MOTHER. One is the evidential evaluation. The other is the practical evaluation which is related to the practical values of believing in her son's innocence. Hookway writes, "[T]he mother has the goal of preserving the reputation of her family," and believing that her son is innocent is valuable for achieving that goal (p. 187). On this practical-consideration account, the akratic belief is possible because the practical consideration exerts a *direct* influence on the belief, and the practical consideration outweighs the evidential consideration.

It is worth emphasising that the practical-consideration account is *not* appealing to a familiar point that our practical consideration may exert an *indirect* influence on our beliefs. One explanation of how such an indirect influence is possible is that our practical consideration can influence the way we collect evidence, and hence influence our overall evidence and evidence-based beliefs.

But this is not what is happening in epistemic akrasia because the akratic subject's overall evidence suffices to reject the belief no matter whether the overall evidence is influenced by the subject's practical consideration.

By contrast, the practical-consideration account claims that the influence of practical considerations is *directly* on the belief. But it is quite controversial to say that practical considerations can exert a *direct* influence on our beliefs. Proponents of the practical-consideration account might argue that they are not making a claim about practical considerations and beliefs in general, rather they are only saying that akratic subjects, in particular, take their practical considerations as a good reason for their beliefs.

But it is far from clear why akratic subjects act in this particular way. Proponents of the practical-consideration account may further argue that akratic subjects are epistemically irrational: an akratic subject may be epistemically irrational in the sense that the subject takes it that whether *p* is true depends not only on the evidence but also on the practical values of believing that *p*; or an akratic subject may be epistemically irrational in the sense that the subject takes it that believing that *p* is believing that *p* is valuable.

This defence, however, can hardly work. This is because epistemic irrationality is a domain-general deficit (Gold & Hohwy, 2000) and akratic subjects do not seem to suffer from a domain-general deficit. Otherwise, they would hold a wide range of bizarre beliefs and would lack a wide range of normal beliefs, both of which are not observed in akratic subjects. The practical-consideration account, therefore, falls short of explaining how the akratic belief is possible.

I think part of Hookway's insight is correct that practical considerations might be a contributory factor in some cases of akratic belief, but appealing to the view that practical considerations exert a *direct* influence on our beliefs can hardly work. The dual-force

framework may offer an alternative view on how practical considerations may contribute to the formation of beliefs. From the perspective of the dual-force framework, practical considerations do not have a direct influence on beliefs, rather they can (directly) influence the way the contents of beliefs are presented to the subjects in their clear experiences. This idea is not completely novel. Mele (2001), for example, writes: “[M]otivation can ... affect the salience of available hypotheses” (pp. 29-30). If we take practical considerations as a source of motivation and take phenomenal clarity as a form of salience, then we can interpret Mele’s words as that practical considerations can affect the phenomenal clarity of available hypotheses. That is to say, the practical values of the hypothesis may make a difference in the way the hypothesis is presented to the subject such that the subject may have a clear experience in which it seems clear to them that the hypothesis is true.

Merely saying that the subject has such a clear experience does not suffice to explain why they fail to reject the hypothesis in the face of the counterevidence. The dual-force account of akratic belief further adds that the subject fails to reject the hypothesis because the primitive causal force of their clear experience may outweigh the causal force of the counterevidence and hence causally compel them to believe what seems very clear to them, despite the fact that they are not justified to believe so.

Admittedly, in some cases the subject *should* not have such a clear experience and the existence of such a clear experience suggests that there might be some deficit in the subject’s experience. We have seen that appealing to a domain-general deficit is problematic for the practical-consideration account. Does this mean that appealing to a deficit in the subject’s experience would be problematic for the dual-force account too? I do not think so. Because a deficit in experience can be a domain-specific deficit (Coltheart, 1999; Fodor, 1983) and hence it accords well with the fact that akratic

subjects do not have a wide range of bizarre beliefs and do not lack a wide range of normal beliefs.

6.2.3. Conflicting belief systems

Consider the following case adapted from Greco (2014):

MATT: (A) Matt believes that flying is dangerous, despite that (B) he also believes that the evidence shows that flying is not dangerous. And (C) Matt is consciously aware of both at the same time.

Matt's belief that flying is dangerous is an akratic belief. But how is it possible? Greco (2014) argues that akratic beliefs are due to a conflict between a linguistic belief system producing beliefs_{*l*} and a non-linguistic belief system producing beliefs_{*n*}. According to Greco, Matt's belief that flying is dangerous is produced by his non-linguistic belief system and he believe_{*n*} that flying is dangerous. The non-linguistic belief system is linked to Matt's emotional response to flying. Greco writes: "Matt is extremely afraid of flying." (p. 202) Furthermore, Greco suggests that Matt also "believes_{*l*} that flying isn't particularly dangerous" which is produced by Matt's linguistic belief system and is based on Matt's evidence (p. 213).

The suggestion that Matt has a belief_{*l*} is controversial. But a more serious problem with Greco's account based on a distinction between belief_{*l*} and belief_{*n*} is that it is hard to see how Greco's account can meet (C) the concurrent consciousness condition of the minimal conception of akratic belief. If we understand Greco's account as that at the time when evidence is not considered, Matt forms and holds a belief_{*n*}, and then at a different time when evidence is considered Matt forms and holds a belief_{*l*}, then Greco's account fails to meet (C). If we understand Greco's account as requiring that Matt is at the same time consciously aware of his belief_{*l*} and belief_{*n*} and the related evidence,

then Greco's account falls short of telling us why Matt fails to reject his belief_n in the light of his evidence, even if we accept that the belief_n is initially produced by a non-linguistic belief system that does not take his evidence into consideration.

In a recent development of Greco's account, Kearl (2019) argues that epistemic akrasia is not explained by a conflict between a non-linguistic belief system and a linguistic belief system, but by a conflict between different "linguistic belief-formation systems, which are perhaps sensitive to different aims, or operative in different contexts." (p. 2514) Unlike Greco's account, Kearl's account does not rely on the view that akratic beliefs are non-linguistic. But, like Greco's account, Kearl's account has difficulties in meeting (C). If the akratic belief and the evidence are processed in different linguistic belief-formation systems, then (C) is not met because different linguistic belief-formation systems are "operative in different contexts"; if the akratic belief and the evidence are processed in the same linguistic belief-formation system, then Kearl's account falls short of telling us why the akratic belief is not responsive to evidence.

I think Greco and Kearl are correct that epistemic akrasia is due to a conflict between different ways of belief formation, but are mistaken in locating the conflict between a linguistic system and a non-linguistic system, or between two linguistic systems. From the perspective of the dual-force account, the akratic belief that p is due to a conflict between the subject's clear experience in which it seems very clear to them that p and the subject's evidence that not-p. The reason why MATT akratically believes that flying is dangerous is that the primitive causal force of his clear experience in which it seems clear to him that flying is dangerous outweighs the causal force of his evidence and hence compels his assent. Unlike Greco's account and Kearl's account, the dual-force account can meet (C) because it allows that the subject's evidence and clear experience are presented in the consciousness at the same time.

In the case of MATT, I think Greco is also correct to link Matt's akratic belief to his emotional response. But Greco is not explicit about how Matt's emotional response could lead to the akratic belief. This unfilled gap between the emotional response and the akratic belief might invite us to wonder whether the link between the emotional response and the akratic belief is real and whether Matt actually holds the akratic belief. Freedman (2017), for example, suggests that Matt does not akratically *believe* that flying is dangerous, rather what seems to be Matt's belief is actually a form of Matt's somatic response which includes Matt's emotional response. This suggestion seems too radical, however. Because Matt does not only have a somatic response to flying but also says that he believes that flying is dangerous. And there is no reason to suppose that Matt suffers from the disability to distinguish between somatic responses and beliefs.

How can the dual-force account help fill in the gap? According to the dual-force account, our emotional responses can have an impact on the phenomenal clarity of our seeming experiences. The more Matt is afraid of flying, the clearer it seems to him that flying is dangerous, and the more he is causally compelled by his seeming experience to believe that flying is dangerous. When Matte is extremely afraid of flying, he may have such a clear experience in which it seems very clear to him that flying is dangerous, whose primitive causal force outweighs the causal force of his evidence and hence causally compels his assent.

6.2.4. Appreciating evidence and being moved by it

In an early discussion of epistemic akrasia, Heil (1984) proposes that there is a "gap between appreciating warrant and coming to hold a belief thereby warranted. ... [I]t is one thing to appreciate evidence, another thing to be *moved* by one's appreciation

of it.” (pp. 69-70, emphasis added) That is to say, for one to actually believe what they believe they ought to believe, they need to be *moved* by their believing that she ought to believe so. According to Heil, an akratic belief arises when the subject fails to be moved by their believing that they ought to believe so.

Heil does not talk much about how the failure is possible. He is more concerned with the question of whether epistemic akrasia, or doxastic incontinence in his terminology, is possible. And for him, “the instances of apparent incontinence” is good enough to establish that epistemic akrasia is possible (p. 69).

When we consider how epistemic akrasia is possible, there seem to be two different ways to develop Heil’s proposal. One is that the akratic subject suffers from a general deficit of being moved by their believing that they ought to believe so. This answer can be discerned in Borgoni and Luthra’s (2017) account of epistemic akrasia. They argue that overcoming the gap between believing so and believing that one ought to believe so requires “the successful exercise of the capacity for critical reasoning” and epistemic akrasia manifests a failure of the subject’s capacity for critical reasoning (p. 885).

However, if the akratic subject suffers from a deficit of a critical reasoning capacity as such, then we would expect to see the subject holding a wide range of bizarre beliefs and lacking a wide range of normal beliefs because a deficit of reasoning capacity is a domain-general deficit (Gold & Hohwy, 2000). This is not what is observed in akratic subjects. Therefore, it is unlikely that akratic subjects suffer from a reasoning deficit.

The second way to develop Heil’s proposal is that akratic subjects commit some performance error. One variant of this answer can also be discerned in various places of Borgoni and Luthra’s (2017) paper. They sometimes seem to suggest that the deficit is not a domain-general reasoning deficit, but a mistake made in the subject’s

exercising her reasoning capacity. But, if this is the case, then the subject should be able to correct the error and give up their akratic belief when they are consciously aware of the conflict between the akratic belief and what they ought to believe, as required by the concurrent consciousness condition (C). But this does not happen in epistemic akrasia. Therefore, it is unlikely that akratic subjects commit some performance error.

I think Heil is correct that in akratic beliefs there is a gap between the subject's appreciation of the evidence and coming to hold the belief that is supported by the evidence. The dual-force framework can offer an alternative explanation of the gap. It argues that whether one would come to believe that *p* can be influenced by both one's appreciation of the evidence related to *p* and one's clear experience in which it seems clear to them that *p*; and the gap is created by the conflict between the akratic subject's evidence and clear experience.

Besides, the dual-force account can remain neutral about whether in ordinary cases there is a general gap between one's appreciation of the evidence and being moved by it; it can also remain neutral about Borgoni and Luthra's (2017) proposal that overcoming the gap requires "the successful exercise of the capacity for critical reasoning" (p. 885). For example, in many ordinary cases in which seeming experience is not a factor that is at odds with the evidence, a version of the dual-force account can agree with Raz (2009) that "[t]here is no gap [contra Heil's proposal], no extra step in reasoning [contra Borgoni and Luthra's proposal], between believing that the case for the truth of the proposition is conclusive and believing the proposition." (p. 39)

6.3. Conclusion

I have argued that the dual-force framework can offer a better explanation of many akratic beliefs. Importantly, the dual-force account of akratic beliefs is not a complete rejection of the current accounts. Instead, the dual-force account can take in many important insights from the current accounts and integrate them into a coherent account: such as the insight on the gap between the appreciation of evidence and being moved by it (Heil, 1984), and the insights on the significance of seeming experience (Scanlon, 1998), salience (Mele, 1986, 2001), practical consideration (Hookway, 2001), and emotion (Greco, 2014) in the aetiology of akratic beliefs.

The explanatory power of the dual-force framework in explaining akratic beliefs, in turn, provides further support to the core proposal of the dual-force framework that beliefs can be understood as the results of the interaction between evidence and clear experience. Taken together, it seems more promising that the dual-force framework can help advance our understanding of delusional beliefs too. In fact, many scholars believe that delusions are the touchstone for evaluating a theory of belief and perception. Stone and Young (1997), for example, write: “[A]ny philosophical theory of belief formation, and of the role of perception in belief formation, must enable a coherent description of ... delusions” (p. 332). If this is true, then we must move to examine how the dual-force framework can help advance our understanding of delusional beliefs, as I will do in the next chapter.

Chapter 4.

The dual-force framework for explaining delusions

[A]nyone who is free from doubt claims he has such 'great light' and has no less strong a propensity of the will to affirm what he has no doubt about than someone who possesses real knowledge. Hence this 'light' can explain why someone obstinately defends or holds on to a given opinion...

(Hobbes, "Thirteenth objection", in Descartes, 1984, p. 134)

1. Introduction

The dual-force framework can advance our understanding of delusions in three respects. First, it can offer a plausible way to flesh out the recurrent personal remarks on the resemblance between delusions and certain non-delusional beliefs, such as that the sky is blue or that $2 + 3 = 5$. Second, it can help make some progress in addressing the evidence and specificity challenges of delusions. Third, it can shed new light on the underpinnings of delusions. This chapter will discuss these three respects in turn. The dual-force framework, however, does not claim to render delusions completely understandable. The last section will highlight some remaining questions for future study.

2. Personal remarks on delusions

At the beginning of the last chapter, we have seen three exemplary remarks on the resemblance between delusions and certain non-delusional beliefs. With the help of the dual-force

framework that is developed on the basis of non-delusional beliefs, we are in a place to return to delusions and say more about what the resemblance is. Let us consider the three exemplary remarks in turn.

When Mithen talks about her own delusions on Twitter, she writes: “They feel like you know something, like you know the sky is blue.”⁸⁰ A parsimonious reading of Mithen’s remark might be that Mithen is merely saying that she is certain of her delusions to the same degree as she is certain that the sky is blue, and nothing more.

This parsimonious reading, however, does not do justice to Mithen’s remark. One obvious reason is that there are a number of different ways we are certain of something, and the parsimonious reading falls short of capturing the distinctive way we are certain that the sky is blue. For example, the way we are certain of the dates of our birthdays is arguably different from the way we are certain that the sky is blue: we are certain of what the dates of our birthdays are because they are what our parents told us or they are what our ID cards say, whereas we are certain that the sky is blue because, when we look at the sky, it seems clear to us that the sky is blue. By comparing her delusions to her knowing that the sky is blue, it is plausible that Mithen’s remark has something more to say than merely stating that she is certain of her delusions.

In contrast to the parsimonious reading, the dual-force framework offers a way to understand the distinctive way we know that the sky is blue: we know that the sky is blue because when we look at the sky, we have a clear experience in which it seems clear to us that the sky is blue. Following this line of reading, Mithen’s remark can be understood as saying that like the way we know that the sky is blue, Mithen thinks she knows that her delusions are true because she has clear experiences in which it seems clear to her that her delusions are true.

⁸⁰ <https://twitter.com/mithenzoe/status/1364426104564449281?s=11>

Similarly, when Kane writes that “[Her husband] ‘knows’ his delusions are true the same way I know I have 5 fingers on each hand”,⁸¹ it can be understood as saying that there is a resemblance between the way her husband thinks he knows that his delusions are true and the way we know that we have 5 fingers on each hand. According to the dual-force framework, we know that we have 5 fingers on each hand because when we look at our hands, we have a clear experience in which it seems clear to us that we have 5 fingers on each hand, and in a similar vein Kane’s husband thinks he knows that his delusions are true because he has clear experiences in which it seems clear to him that his delusions are true.

The experience in which it seems clear to us that the sky is blue and the experience in which it seems clear to us that we have 5 fingers on each hand are both clear sensory experiences. Compared with the talk of the resemblance between delusions and beliefs that are based on clear sensory experiences, the talk of the resemblance between delusions and beliefs that are based on clear intellectual experiences is relatively hard to find. This may be because in our daily lives the talk of intellectual experiences itself is not as frequent as the talk of sensory experiences. Nonetheless, when a comparison between delusions and beliefs that are based on clear intellectual experiences is raised, people with delusions do seem to emphasise that there is an important resemblance between them. For example, consider the following report in which Nash, the Nobel Laureate, was asked by Mackey about his delusions:

“How could you,” began Mackey, “how could you, a mathematician, a man devoted to reason and logical proof ... how could you believe that extraterrestrials are sending you messages? How could you believe

⁸¹ https://twitter.com/syl_kane/status/1364461903767097344?s=11

that you are being recruited by aliens from outer space to save the world? How could you ... ?”

Nash looked up at last and fixed Mackey with an unblinking stare as cool and dispassionate as that of any bird or snake. “Because,” Nash said slowly in his soft, reasonable southern drawl, as if talking to himself, “the ideas I had about supernatural beings came to me the same way that my mathematical ideas did. So I took them seriously.” (Nasar, 1998, p. 11)

What did Nash mean by “the same way”? One interpretation is that Nash was saying that his delusions were reached via a “logical reasoning” process (e.g. Ashinoff et al., 2021, p. 15). I find it doubtful that this is what Nash meant. Not only is it an unintelligible interpretation that Nash was saying that there were logical reasons that supported his delusions in the same way his mathematical proof supported his mathematical ideas, but there is also no evidence that here Nash was talking about logical reasons at all. Otherwise, Nash might have provided the specific logical reasons for his delusions and explained why those logical reasons supported his delusions in the same way his mathematical proof supported his mathematical ideas.

An alternative, and more plausible, interpretation is that the purpose of talking about mathematical ideas in his remarks on delusions was not to talk about how logical reasons supported mathematical ideas, but to talk about the distinctive phenomenal way certain mathematical ideas struck us (e.g. Rey, 2020). It is a well-known point that certain mathematical ideas can strike us to be true, in the absence of logical reasons proving that they are true. As Gödel (1984) vividly puts it, “[T]he axioms force themselves upon us as being true.” Furthermore, Gödel also emphasises the similarity between the force of intellectual experience and the force of sensory experience. He writes: “I don't see any reason why we should have less

confidence in this kind of perception, i.e., in mathematical intuition, than in sense perception” (p. 484). Following this line of thought and in the terminology of the dual-force framework, Nash was saying that the contents of his delusions struck him in the same way his mathematical ideas did: in both cases he had clear experiences in which it seemed clear to him that the contents were true; and it was the clear experiences, rather than some logical reasons, that compelled his assent.

It is worth noting that the dual-force account does not claim that all delusions must strike their subjects with the same degree of phenomenal clarity: it does not claim that all delusions must strike their subjects in the same way we perceive that $2 + 3 = 5$ or that we have 5 fingers on each hand. Instead, it only proposes that in many cases of delusions the subjects may have a clear experience in which it seems clear to them that the contents of their delusions are true, and the degree of the phenomenal clarity may vary from case to case.

To sum up, by highlighting the close resemblance between delusions and beliefs that are based on clear sensory/intellectual experiences, the dual-force framework offers a plausible interpretation of the recurrent remarks on the resemblance between delusions and certain non-delusional beliefs. Needless to say, there may be other interpretations of these remarks and people with delusions do make other sorts of remarks. In the end, whether a particular interpretation should be taken seriously depends on whether it can help explain the bafflement revolving around delusions. In the next section, I shall argue that the dual-force framework can help make some progress in addressing the two central challenges of delusions: the evidence challenge and the specificity challenge.

3. The dual-force account of delusions

3.1. Clear experiences and delusions

In Chapter 2, I have argued that current theories, which appeal to anomalous experiences and/or reasoning abnormalities to explain delusions, have difficulties in addressing the evidence and specificity challenges. As regards the Capgras delusions, notable examples of the proposed anomalous experience⁸² include the experience of reduced affective responses or the experience of reduced familiarity, and the experience that “There’s something odd about this woman.” (Section 2 of Chapter 2); notable examples of the proposed reasoning abnormalities include cognitive biases (e.g. the JTC bias), and the impairment of working memory and/or executive function (Section 3 of Chapter 2). While the specificity of the anomalous experiences can help explain why the subject only has the Capgras delusion rather than a wide range of delusional or odd beliefs, the anomalous experiences fall short of explaining why the subject fails to reject the delusion in the face of the significant counterevidence: the delusion is implausible in the light of common knowledge; it is at odds with the fact that the “imposter” looks like his wife and knows things that only his wife knows; it is also at odds with the testimonies from his friends

⁸² I use the term *anomalous experience* in the narrow sense that refers to the kind of experience which is proposed as an antecedent causal factor by current theories of delusions, particularly Maher’s one-factor theory, the two-factor theory and the salience theory. In this narrow sense, patients’ anomalous experiences are distinct from their abnormal experiences of having delusions, which are the focus of the phenomenological approach (see e.g. Sass & Pienkos, 2013); patients’ anomalous experiences are also distinct from the clear experiences, proposed by this present chapter, in which it seems clear to them that the contents of their delusions are true. By contrast, one might take the term *anomalous experience* in the broad sense, according to which to say patients have anomalous experiences is just to say that they have some experiences that people without delusions do not have. In this broad sense, all theories of delusions are talking about anomalous experiences. This way of talking of patients’ experiences, however, appears to be uninformative. I do not use the term *anomalous experience* in this broad sense.

and relatives; and the subject may even have partial insights into the force of the counterevidence and the implausibility of the delusion. As for reasoning abnormalities, they can help explain why the subject fails to reject the delusion in the face of the significant counterevidence: it is because reasoning abnormalities prevent the subject from properly evaluating the delusion in the face of the counterevidence. But, as a form of domain-general factor, reasoning abnormalities are at odds with the specificity of the Capgras delusion. The two-factor theory, which takes in both anomalous experiences and reasoning abnormalities, makes significant improvements but also inherits some of the shortcomings of both.

Why is it so challenging for current theories to explain delusions? I think the deeper reason may be that current theories assume a simplistic evidential-force framework of belief, according to which the only two factors in the aetiology of a subject's belief are the related evidence and the subject's ability to evaluate a hypothesis in the light of the evidence. From the perspective of this simplistic framework, the anomalous experience is the subject's evidence for the delusion, and reasoning abnormalities take a hold of the subject's ability to properly evaluate the delusion in the light of the counterevidence. That is to say, despite the anomalous experiences and reasoning abnormalities, delusions are, by and large, evidence-based beliefs. The problem with thinking of delusions as evidence-based beliefs is that there appears no way to satisfactorily solve the tension between the subject's flouting of the evidence related to their delusions, on the one hand, and the subject's intact ability to evaluate the evidence related to their non-delusional beliefs, on the other hand.

Of course, evidence and one's ability to evaluate evidence are important factors in the formation and maintenance of beliefs. But they may not be the only factors. From the perspective of the dual-force framework, the subject's clear experience in which it seems

clear to them that p can also exert an impact on whether the subject believes that p .

It is a mundane fact that when we have a clear experience in which it seems clear to us that p , the clear experience may incline us to believe that p . When we look at the sky and it seems clear to us that the sky is blue, this clear sensory experience may incline us to believe that the sky is blue; when we look at our hands and it seems clear to us that we have 5 fingers on each hand, this clear sensory experience may incline us to believe that we have 5 fingers on each hand; when we consider the proposition that $1 + 1 = 2$ and it seems clear to us that $1 + 1 = 2$, this clear intellectual experience may incline us to believe that $1 + 1 = 2$. When we are taking our morning walk and it seems clear to us that there is a fox hiding in the woods, this clear experience may incline us to believe that there is a fox hiding in the woods. In cases like these, we may or may not have possessed independent evidence for our beliefs, and we may or may not continue to search for independent evidence to verify our beliefs, but, in any case, the basic point appears plausible that clear experiences can be contributory factors in many of our beliefs.

Can a clear experience be a contributory factor in the aetiology of the Capgras delusion? Let us start by considering ordinary non-delusional cases. Most people do not have the Capgras delusion, but it is arguable that we do occasionally have the experience in which it seems to us that a person is not who they look like. For example, Coltheart et al. (2010) suggest that the prior probability of the hypothesis that a woman who looks like the subject's wife is not his wife is 0.01, whereas McKay (2012) suggests that the prior probability is 0.00027. While they are talking about the probability about whether the woman is in fact not his wife, rather than about the probability of the clear experience in which it seems to the subject that the woman who looks like his wife is not his wife, their points lend support to the view that every time an ordinary person looks at his wife, it is possible

for them to have a clear experience in which it seems to them that the woman who looks like his wife is not his wife.

Admittedly, for an ordinary person, such a clear experience would not suffice to make him believe that his wife is an imposter after he recognises the abundant counterevidence. But it does exert an impact on the person's belief in the sense that at first sight it may causally incline the person to believe what he experiences, and, according to phenomenal dogmatists, the experience may also offer *prima facie* justification. Put it another way, even though in ordinary cases the causal force and justificatory force of the person's clear experience are eventually outweighed by the causal force and justificatory force of the counterevidence, his belief that this woman is indeed his wife is not a purely evidence-based belief, but the result of the interaction between his clear experience and his evidence.

Similarly, the Capgras delusion may also be understood as the result of the interaction between the patient's clear experience and the evidence. What distinguishes the patient's clear experience from an ordinary person's clear experience is the degree of the phenomenal clarity of the experience: the patient's experience may have a higher degree of phenomenal clarity with a much stronger causal force. This distinction between the clear experiences in delusional cases and ordinary non-delusional cases is also broadly in line with the self-reports of some patients with delusions. For example, in the interview by Feyaerts et al. (2021b), a patient suffering from delusions said: "It is really a more compelling, a much-too-compelling reality. Much more compelling than ordinary reality." (p. 792) It is worth noting that this patient was diagnosed with schizophrenia and appeared to have polythematic delusions. It remains to be shown whether patients with monothematic delusions make similar reports. The dual-force account of delusions predicts that we may find similar reports by patients with monothematic delusions when they are properly interviewed.

Suppose that a patient does have a *compelling* clear experience in which it seems clear to them that the content of the delusion is true. The immediate questions are: how compelling can the patient's clear experience be?; and, more specifically, how can the patient's clear experience outweigh the significant evidence that the content of the delusion is false? To answer these questions, we need to take a closer look at the interaction between the causal force and justificatory force of the patient's clear experience and the causal force and justificatory force of the counterevidence. According to the dual-force framework developed in Chapter 3, a set of plausible answers may be as follows.

When the patient has a clear experience in which it seems clear to him that his wife is an imposter, (following phenomenal dogmatists) the clear experience may offer limited *prima facie* justification for believing that his wife is an imposter. But this limited *prima facie* justification is defeasible and should be defeated in the light of the significant counterevidence: the delusion is implausible in the light of common knowledge; it is at odds with the fact that the "imposter" looks like his wife and knows things that only his wife knows; it is also at odds with the testimonies from his friends and relatives; and the subject may even have partial insights into the force of the counterevidence and the implausibility of the delusion. Therefore, the delusional belief is not justified (see also Section 4.2 of Chapter 3).

Nonetheless, it is possible that the causal force of the clear experience that his wife is an imposter may outweigh the causal force of the counterevidence, and hence may causally compel the patient to believe that his wife is an imposter. This possibility is due to the asymmetries between how the clear experience, on the one hand, and the counterevidence, on the other hand, causally affect the patient's belief. As I have argued in Chapter 3, the causal force of clear experience is primitive and independent of its justificatory force. As the phenomenal clarity of the experience increases, its causal force

may keep increasing, even when its justificatory force of the experience is outweighed by the justificatory force of the counterevidence. By contrast, the causal force of evidence is neither primitive nor unlimited: it may reach its peak when the evidence is sufficient. These asymmetries make it possible that there are cases in which a person has a very clear experience with enormous causal force but limited justificatory force, on the one hand, and lots of counterevidence with enormous justificatory force but only limited causal force. In such cases, even though the justificatory force of the counterevidence outweighs the justificatory force of the clear experience, and the person hence is not justified to believe what seems very clear to them; but the causal force of the clear experience may outweigh the causal force of the counterevidence, and the person hence is causally compelled to believe what seems very clear to them. The patient with the Capgras delusion, I propose, may be in such a case. Put simply, it is the primitive causal force of the patient's clear experience that helps explain why he is causally compelled to believe that the content of his delusion is true, even in the light of the significant counterevidence.

When we discuss the explanatory power of reasoning abnormalities, the main pitfall is that the domain-general nature of reasoning abnormalities is at odds with the specificity of delusions. Does clear experience, or in particular its phenomenal clarity, have a similar pitfall?

Given that there have not been (systematic) empirical studies on phenomenal clarity and its impairment, I do not have conclusive empirical evidence to report. But, on the face of it, since the abnormal phenomenal clarity is an abnormality in the patient's subjective experience underpinned by domain-specific modular systems (Coltheart, 1999; Fodor, 1983), it is likely that the abnormal phenomenal clarity and its underpinning are relatively restricted. Therefore, compared with domain-general reasoning abnormalities,

the notion of the abnormal phenomenal clarity accords better with the specificity of delusions.

I suspect that the main motivation for thinking that the phenomenal clarity in delusions is a global abnormality rather than restricted to certain specific experiences is that, at first sight, the notion of clear experience might be taken as the same as what has been called hyper-reality experience, and, according to some phenomenological theories, patients with schizophrenia often have *global* hyper-reality experiences (see e.g. Feyaerts et al., 2021a; Feyaerts et al., 2021b; Ratcliffe, 2017; Van Duppen, 2015).

However, I think it is a mistake to equate clear experience with hyper-reality experience, or to equate phenomenal clarity with hyper-reality. Admittedly, when a patient has a clear experience in which it seems clear to them that the content of their delusion is true, the phenomenal clarity of their experience can be taken as a reason why they think that their delusion is real. In this sense, there might be a connection between phenomenal clarity and hyper-reality.⁸³ But a closer look would reveal that there are significant differences between them.

The term *hyper-reality* is used to refer to a wide range of heterogeneous phenomena. For example, consider the following words by patients with delusions, which Feyaerts et al. (2021b) use to illustrate what they mean by hyper-reality:

⁸³ In my view, the proposal that clear experiences are important factors in the aetiology of delusions is *broadly* in line with phenomenological theories' emphasis on the significance of altered reality experience, although, strictly speaking, neither my notions of clear experience and phenomenal clarity nor the literature on Cartesian clarity and phenomenal dogmatism, based on which the notions are developed, have been discussed by current phenomenological theorists. For a brief discussion of the relationship between my notion of phenomenal clarity and Jaspers' view on reality experience, see the beginning of Chapter 3. It is also worth noting that some recent phenomenological theorists occasionally present their views as incompatible with thinking of delusions as beliefs, and as incompatible with current cognitive theories of delusions (see e.g. Feyaerts et al., 2021a; Feyaerts et al., 2021b). By contrast, from the perspective of the dual-force framework, clear experiences not only are compatible but also help us explain why delusions are beliefs (see Appendix). I take clear experiences as complementing rather than replacing the factors proposed by current cognitive theories.

“During a psychosis, it is all so intense, all so utterly lifelike. You know that everything is meant to be.”
(Bert)

“I had a sort of heightened perception; I saw connections everywhere, connections which I alone saw, for example, on the doors of the psychiatric ward. The semantics of words revealed a hidden meaning.”
(Kurt) (p. 789)

In their summary, Feyaerts et al. (2021b) write: “things and events in hyper-reality seemed to be permeated by an overall sense of necessity, compulsion, and heightened meaningfulness.” (p. 791) By contrast, in my understanding of the patients’ words, phenomenal clarity might or might not be part of what the patient Bert meant by “lifelike”; but, unlike hyper-reality, phenomenal clarity does not refer to the patients’ other abnormal experiences including the feeling that everything is meant to be, the perceived promiscuous connections, or the hidden meaning. Given the significant differences between the phenomena that phenomenal clarity aims to capture and the phenomena that hyper-reality talks about, it is not straightforward that the abnormal phenomenal clarity is a global deficit, even if it is true that in schizophrenia the hyper-reality is a global deficit.

It might pose a challenge to the dual-force account if there were some cases in which the patients have globally abnormal clear experiences in which everything, including a wide range of bizarre or odd contents, seems clear to them to be true, on the one hand, but they only have a few delusions, on the other hand. To my best knowledge, no such cases have been reported. On the contrary, in the reported cases when patients talk about how something is abnormally real, they often talk about the experiences that are directly related to their delusions. This suggests that there is a covariation between their clear experiences and their delusions. If this is true, then,

unlike reasoning abnormalities, clear experiences can help us address both the evidence and specificity challenges of delusions.

3.2. Current theories and the dual-force framework

It is, however, important to re-emphasise that the dual-force framework is not proposing that delusions are purely seeming/clear experience-based beliefs, but merely that patients' clear experiences are one of the contributory factors in the aetiology of delusions. The dual-force account of delusions can accept that the anomalous experiences and reasoning abnormalities proposed by current theories may also play contributory roles, though it is a remaining task for current theories to explain how these factors can tackle the evidence and specificity challenges.

There is no mystery in saying that one's clear experience that *p*, one's evidence related to *p*, and one's ability to evaluate evidence all play important roles in the aetiology of one's belief that *p*. To see the point, consider a scenario adapted from Austin's pig case, in which the subject may see the pig-like marks and the pig food which count as their evidence, know that the evidence supports that there is a pig, and also have a clear experience in which it seems clear to them that there is a pig; in such a scenario, the subject's belief that there is a pig may be based on all these factors. By analogy, in the Capgras delusion, the subject may have an anomalous experience with reduced affective responses to his wife, which counts as his evidence for the delusional hypothesis that his wife is an imposter, suffer from some reasoning abnormality which prevents him from properly evaluating the hypothesis in the face of the counterevidence, and also have a very clear experience in which it seems very clear to him that his wife is an imposter; and all these factors, the dual-force framework proposes, may play contributory roles in the aetiology of the Capgras delusion.

It is also likely that there are mutual interactions between the subject's clear experience and the anomalous experience/reasoning abnormality. It is a familiar point in the literature that clear experience can be affected by many factors including background knowledge, desires, and unjustified beliefs (see e.g. Tucker, 2013a). In Section 6 of Chapter 3, I have argued that one's practical considerations and emotions may also affect their clear experiences. Here, the subject's anomalous experience and his reasoning abnormality may also play a contributory role in the aetiology of the subject's clear experience in which it seems very clear that his wife is an imposter. But it is worth noting that they are by no means the only contributory factors, because in the light of the counterevidence as well as the subject's anomalous experience and reasoning abnormality, I think, it should not seem very clear to the subject that his wife is an imposter. The fact that it does seem very clear to the subject that his wife is an imposter suggests that the abnormally clear experience itself is a relatively independent factor in the aetiology of the delusion. Conversely, the subject's clear experience may also have some impact on his affective responses to his wife and the way he evaluates the counterevidence: if it seems very clear to the subject that his wife is an imposter, this clear experience may be a reason why he has reduced affective responses to his wife and why he does not fully appreciate the force of the counterevidence. In the light of the dual-force framework, we may naturally wonder how exactly these factors interact with each other and how much contribution a particular factor makes in the aetiology of a particular delusion. I take it that they are empirical questions, about which the answers may vary from case to case.

In short, the core proposal of the dual-force framework is quite minimal: that is, the subject's delusion can be understood as the result of the interaction between the evidential force of the subject's evidence, including both the anomalous experience in favour of the

delusion and the significant counterevidence, and the phenomenal force of the subject's clear experience. Compared with current theories, it adds that clear experiences are contributory causal factors in the aetiology of delusions.

With this minimal proposal, the dual-force framework aims to enrich rather than replace current theories of delusions. In the rest of this section, I shall explain in more detail about how the dual-force framework can help the two-factor theory as well as some other theories to explain delusions.

3.2.1. The two-factor theory and the dual-force framework

3.2.1.1. The endorsement account of delusion adoption

The two-factor theory distinguishes between two accounts of how a delusion is initially adopted: the endorsement account and the explanationist account (Davies & Egan, 2013). According to the endorsement account, the content of the delusion is already encoded in the anomalous experience, and the delusion arises as the endorsement of the content of the anomalous experience. The subject with the Capgras delusion, for example, may have an anomalous experience that "This woman looks like my wife but she is not my wife", and the Capgras delusion is adopted as the endorsement of this anomalous content. The endorsement account, however, has difficulties in addressing the evidence challenge. (Section 3.3.3 of Chapter 2)

Regarding the endorsement account, the dual-force framework takes it that the content of the delusion being encoded in the anomalous experience itself may be the subject's evidence for his delusion, but this evidence is far from sufficient to outweigh the significant evidence against his delusion. To explain the adoption of the delusion, the dual-force framework suggests that the

consideration of the evidence is not the only factor in the aetiology of delusions, the other factor may be the abnormal phenomenal clarity of his experience, whereby it seems clear to him that the content of his experience is true. That is to say, there are two kinds of anomalies in the subject's abnormal experience: one is the anomalous content that "This woman looks like my wife but she is not my wife"; the other is the abnormal phenomenal clarity. In this sense, the subject's experience is a form of abnormal clear experience. When the subject endorses the content of his experience, the subject is not simply exercising the "default or prepotent doxastic response" to his experience (Davies & Egan, 2013, p. 707), but is *causally compelled* by his abnormal clear experience to endorse what seems very clear to them.

3.2.1.2. The explanationist account of delusion adoption

According to the explanationist account of delusion adoption, the content of the delusion is not encoded in the anomalous experience and the delusion arises as the explanation of the anomalous experience. In the Capgras delusion, the subject's anomalous experience may be the experience of reduced affective responses, or the experience of reduced familiarity, or the experience that "There's something odd about this woman." And the Capgras delusion is adopted by the subject as the explanation of his anomalous experience.

The explanationist account, however, has difficulties in explaining why the delusional hypothesis is adopted in the light of the counterevidence and other more plausible hypotheses which can better explain the anomalous experience: such as the brain-damage hypothesis and the fading-love hypothesis suggested by Coltheart (2007, p. 1059). In other words, there is an explanatory gap in the explanationist account between the anomalous experience as the

explanandum and the delusional hypothesis as the explanans. (Section 3.3.2 of Chapter 2)

Here it is useful to note that the explanationist account is supposed to answer two distinct questions at the same time. One is the content question: that is, why does the delusion have such content, as opposed to some more plausible content? The answer is that it is the subject's explanation of the anomalous experience. The other is the attitude question: that is, why does the subject *believe* so? The answer is that it is the subject's explanation of the anomalous experience. The explanatory gap in the explanationist account, in turn, indicates that it has difficulties in answering both questions.

Regarding the explanationist account, the dual-force framework takes it that the anomalous experience as the explanandum of the delusion may be the subject's evidence in favour of his delusion, but such evidence is far from sufficient to outweigh the significant evidence against his delusion. How can the dual-force framework help the explanationist account explain the adoption of the delusion? Can it appeal to a similar suggestion that the content of the delusion is presented to the subject in a phenomenally clear way in his clear experience? There might be some worries.

Recall that a main disparity between the endorsement account and the explanationist account is whether the content of delusion is encoded in the subject's anomalous experience. In my understanding, the kind of experience at issue here is sensory experience. That is to say, when the dual-force framework suggests that the content of the delusion is presented to the subject in a phenomenally clear way, this suggestion can be read as that the content of the delusion is presented in the subject's sensory experience in a phenomenally clear way. Given that in the case of the explanationist account the content of the delusion is not encoded in the subject's sensory experience, one might wonder whether the dual-force framework can

appeal to a similar suggestion that the content of the delusion is presented to the subject in a phenomenally clear way.

I think it can, because phenomenal clarity can be present in both sensory experiences and intellectual experiences. Therefore, even if the content of the delusion is not encoded in the subject's sensory experience, it is still possible that the content is presented in the subject's intellectual experience in a phenomenally clear way. This explanation is also in line with Nash's remark on the resemblance between his delusions and certain mathematical ideas (Nasar, 1998, p. 11). In this sense, the dual-force framework can help the explanationist account answer the attitude question: the subject adopts the delusion in the face of the counterevidence partly because the subject has a clear intellectual experience whose primitive causal force outweighs the causal force of the counterevidence.

Admittedly, by saying that the subject has a clear intellectual experience in which it seems clear to them that the content of the experience is true, the dual-force framework does not tell us why it is this content, rather than some more plausible content, that is presented to the subject in a phenomenally clear way. Regarding the Capgras delusion, it does not tell us why it is the imposter hypothesis, rather than the brain-damage hypothesis or the fading-love hypothesis, that is presented to the subject in a phenomenally clear way. Therefore, it does not contain an answer to the content question (see also Parrott, 2016). But this only means that more work is needed to understand the aetiology of the subject's clear experience, and it should not be a reason to deny that the subject's clear experience may be a contributory factor in the aetiology of delusions.

3.2.1.3. The persistence of delusions

Once the delusion is adopted, why does the subject not reject it in the face of the counterevidence? According to the two-factor

theory, it is because the subject suffers from reasoning abnormalities. More specifically, the subject suffers from an impairment of executive function and/or working memory which prevents them from properly evaluating the delusion in the light of the counterevidence. Since the impairment of executive function and/or working memory is a domain-general deficit, it has difficulties in explaining the specificity of delusions: that is, why does the domain-general deficit not result in a wide range of delusional or odd beliefs but only result in delusions with specific themes? (Section 3.3.4 of Chapter 2)

Regarding reasoning abnormalities, the dual-force framework can accept that certain reasoning abnormalities may play a contributory role in the aetiology of delusions, but the difficulties of reasoning abnormalities in explaining the specificity of delusions caution us against overemphasising the contribution of reasoning abnormalities. It is unlikely that reasoning abnormalities are the main or the only factor responsible for the persistence of the delusion. From the perspective of the dual-force framework, the subject may have a domain-specific clear experience, in which it seems clear to them that the content of the delusion is true; and it is the compelling causal force of the clear experience that may help maintain the delusion in the light of the counterevidence.

3.2.1.4. Some clarifications

Through the adoption and persistence stages, I have appealed to the idea that the subject may have a clear experience whose primitive causal force compels their assent, even in the light of the counterevidence. One might worry whether it is possible for the clear experience to compel the subject's assent when the subject also has clear experiences in which it seems clear to them that there is abundant counterevidence.

I think it is possible, because phenomenal clarity comes in degrees, and there is no reason to stipulate that it is impossible for the content of the delusion to have a higher degree of phenomenal clarity than the counterevidence. This point is vividly illustrated by the following example:

On admission, [the patient who is a 48-year old musician] repeated that nothing was true: “The hospitals do not exist, the doctors do not exist. . .” He appeared perplexed and frightened, looking at his hands and touching his face repeatedly, and stated that nothing existed, including himself. . . . During the initial 24 hours, the patient refused to eat. He continuously stared at his hands, sometimes stating, “I do not have hands,” “Nothing exists,” . . .(Ramirez-Bermudez et al., 2010, p. 411)

What was happening in this patient with the Cotard delusion? Clearly perceiving our hands is arguably one of the paradigmatic clear sensory experiences (Moore, 1939/2013). The patient appeared to be able to appreciate, to a certain extent, the force of the evidence that his delusions were not true. This was why he repeatedly looked at his hands and touched his face. But why did he not reject his delusions in the face of the counterevidence? From the perspective of the dual-force framework, the subject’s clear experiences that the contents of his delusions were true came with a higher degree of phenomenal clarity, whose causal force outweighed the causal force of his experience of the counterevidence and hence compelled his assent.

So far, I have explained how, by adding a third factor, i.e. the abnormal phenomenal clarity of the patient’s sensory/intellectual experience, the dual-force framework can help the two-factor theory explain delusions. This, however, does not mean that the dual-force framework is committed to a three-factor theory of delusions. The dual-force framework may allow the possibility that in some cases the

patient has no reasoning abnormality at all, and apart from the first factor that explains the content of the delusion, the only other factor is the abnormal phenomenal clarity of the patient's experience. In such cases, the dual-force framework can be taken as offering a new version of the two-factor theory.

3.2.2. Other theories of delusions and the dual-force framework

The two-factor theory is arguably "the most influential neurocognitive account of delusion in the scientific literature"(Braun & Suffren, 2011, p. 2). But it is worth mentioning that the dual-force framework is a generic framework of belief, which can be compatible with other theories of delusions as well.

The dual-force framework can be compatible with Maher's one-factor theory. According to Maher, delusions arise as normal explanations of anomalous experiences. In Section 2 of Chapter 2, I have argued that despite a series of arguments against Maher's one-factor theory, the theoretical possibility is not excluded that there might exist a certain kind of anomalous experience of which the best explanation our normal cognitive system can offer is the delusional hypothesis.

Here, the dual-force framework can add the point that the delusional hypothesis may also be presented to the subject in a phenomenally clear way such that the subject has a clear experience whose primitive causal force compels assent. The clear experience may or may not be a redundant causal factor, depending on whether the delusional hypothesis being the explanation of the anomalous experience is sufficient to outweigh the evidential force of the counterevidence. I shall leave the answer to the latter question open.

The view that clear experience causally compels assent may also be helpful for Currie's imagining theory of delusions, according to which delusions initially arise as imaginings through a process of

imagination, and then these imaginings are misidentified as beliefs by the subjects (Currie, 2000; Currie & Jureidini, 2001; Currie & Ravenscroft, 2002). The imagining theory has difficulties in explaining why the subjects misidentify imaginings as beliefs. One explanation is that subjects with delusions have the disability of distinguishing imaginings from beliefs. But subjects with delusions do not seem to suffer from a general disability of distinguishing imaginings from beliefs, otherwise they would have a wide range of delusional or odd beliefs that are related to imaginings (see also Bayne & Pacherie, 2005).

How could the causal force of clear experience help? First, given that phenomenal clarity can be present in an imaginative experience, it is possible that the content of the imagining may be presented in a phenomenally clear way such that the clear imaginative experience compels the subject's assent. Second, since clear experience is a form of subjective experience whose abnormality, i.e. its abnormal phenomenal clarity, can often be domain-specific (Coltheart, 1999; Fodor, 1983), it accords well with the fact that subjects with delusions do not have a domain-general disability of distinguishing imaginings from beliefs.

The combination of the causal force of clear experience and Currie's imagining theory also offers an intuitive reading of Locke's view on delusions. Locke (1690/1975) wrote:

For by the violence of their Imaginations, having taken their Fancies for Realities, they make right deductions from them. Thus you shall find a distracted Man fancying himself a King, with a right inference, require suitable Attendance, Respect, and Obedience: Others who have thought themselves made of Glass, have used the caution necessary to preserve such brittle Bodies. (p. 161, emphasis added).

Here the violence of imaginations can be read as the brute causal force of the subjects' clear imaginative experiences that compel assent.

The broad compatibility of the dual-force framework, in particular the view that clear experience causally compels assent, understandably, invite the concern about whether saying that the content of the delusion is presented to the subject in a phenomenally clear way or saying that the subject has a clear experience in which it seems clear that the content of the delusion is true is nothing but a redescription of the fact that the subject has a delusional belief: if it were merely a redescription, then no wonder it is compatible with so many different theories of delusions.

I think it is a mistake to take clear experience to be the same as belief. One main reason is that they are dissociable phenomena. On the one hand, there are many illusory and hallucinatory cases in which we have a relatively clear experience that *p* but without believing that *p*. On the other hand, there are also many cases in which we believe that *p* in spite of the fact that it does not seem clear to us that *p*. For example, we believe that the mathematical theories published in prestigious journals are true. But rarely is it the case that when we consider these mathematical theories, it seems clear to us that they are true. More often than not, they seem obscure to us. In such cases, our beliefs are based on the evidence, e.g. the fact that they are published in prestigious journals, rather than based on clear experiences. (For more about the differences between clear experiences and other mental phenomena, see Section 3 of Chapter 3)

Another concern about the value of the notion of clear experiences is how much it can tell us about the underpinnings of delusions. So far, our discussion that clear experiences may be a contributory factor in the aetiology of delusions has focused on the personal level (Davies, 2000a, 2000b; Dennett, 1986). The

subpersonal-level underpinning of phenomenal clarity has not been explored. By contrast, both anomalous experiences and reasoning abnormalities have relatively well-defined subpersonal-level underpinnings. For example, according to the two-factor theory of the Capgras delusion, the subject's anomalous experience is related to the reduced activity in the autonomic nervous system and the reasoning abnormality is empirically related to an impairment of the right lateral prefrontal cortex (for reviews, see Aimola Davies & Davies, 2009; Coltheart et al., 2018). For the dual-force account of delusions to go beyond being a philosophical conjecture (as some neuroscientists might complain) to being an empirically more informative theory, it is essential to fill in this blank. In the next section, I shall discuss a potential candidate for the underpinning of phenomenal clarity.

4. Phenomenal clarity, salience, and dopamine

The most straightforward way to discuss the underpinning of a factor in delusions is to focus on the data of the experiments that directly measure that factor (see e.g. Darby et al., 2017). This strategy, however, has not been feasible for the discussion of phenomenal clarity because, to my best knowledge, there is no experiment that directly measures phenomenal clarity. In the following, I shall take an alternative strategy. I shall focus on a prevalent neurobiological theory of delusions, i.e. the salience theory, in particular the variant called the predictive processing theory,⁸⁴ and argue that there is good reason to think that the underpinning of phenomenal clarity may be similar to the underpinning of salience proposed by the salience theory of delusions.

⁸⁴ For such an interpretation of the predictive processing theory, see e.g. Howes et al. (2020) which is co-authored with Corlett, who is a main proponent of the predictive processing theory.

There are many similarities between the predictive processing theory and the explanationist account of the two-factor theory. Both have been cast in Bayesian terms.⁸⁵ In Bayesian inference, the posterior probability of a hypothesis depends on both its prior probability and its likelihood on the evidence: i.e. how well it explains the evidence including the anomalous experience. In the Capgras delusion, it is clear that the wife hypothesis has a higher prior probability than the imposter hypothesis; but the imposter hypothesis better explains the anomalous experience, i.e. the imposter hypothesis has a higher likelihood than the wife hypothesis when the anomalous experience is concerned. The question is: why does the imposter hypothesis have a higher posterior probability? One version of the two-factor theory argues that overall the ratio of the likelihoods outweighs the ratio of the prior probabilities (Coltheart et al., 2010); another version adds that a bias towards explanatory adequacy is also responsible for the outweighing (McKay, 2012). Despite the disagreement on whether a bias is involved, both agree that the imposter hypothesis rather than the wife hypothesis is adopted because the ratio of the likelihoods outweighs the ratio of the prior probabilities.

In the same vein, proponents of the predictive processing theory write:

In our Bayesian, predictive learning scheme, Capgras [delusion] results when patients experience an anomalous lack of affective responding [which is related to the reduced skin conductance responses] when confronted with their [spouses or] relatives (Ellis

⁸⁵ I will present my argument with the help of Bayesian inference because it is widely used to model the computation of the mind in the literature. However, I take it that the conclusion that some malfunctioning salience-attribution mechanism may underpin the phenomenal clarity of the contents of delusions is more general in that it is compatible with other models of the mind (for reviews of the models of the mind, see Colombo et al., 2021; Genin & Huber, 2021; Halpern, 2017).

& Young, 1990), the delusion constitutes a new ...[belief] driven by the experience, a means for explaining it away (Young, 2008). (Corlett et al., 2010b, p. 360)

Like the Bayesian versions of the two-factor theory, the predictive processing theory argues that the delusion is an explanation of the anomalous experience which is related to the reduced activity in the autonomic nervous system. More specifically, both agree that when the subject sees his wife's face, he expects to have a high-level autonomic activity and the mismatch between the *expected* high-level autonomic activity and the *received* reduced autonomic activity generates a salient prediction error signal (e.g. Coltheart, 2005; Corlett et al., 2010b; Davies & Egan, 2013). And this salient prediction error signal "demands" explanation (Corlett et al., 2010b).

Similarities aside, the predictive processing theory has made distinctive contributions. First, it puts more emphasis on the difference between the prediction error signal itself and its salience. The notion of salience refers to how much the prediction error signal is weighted in the Bayesian inference. The more weighted [i.e. more salient] the prediction error signal is, the more it demands explanation and the less likely it is ignored as noise.⁸⁶ Second, it proposes that, at the neurobiological level, the prediction error signal and its salience are modulated by various neuromodulators, among which dopamine is the most well-known. These two points are evident in the following paragraph by Corlett et al. (2010b):

A crucial distinction ... is between prediction errors per se and the precision or uncertainty [i.e. salience] about those errors. ... This distinction is important

⁸⁶ Apart from being referred to as salience, the weighting of the prediction error signal is also referred to as "uncertainty, novelty or precision" (Corlett et al., 2010b, p. 347).

because it is easy to confuse the role of phasic dopaminergic discharges as encoding reward prediction error ... and the role of dopamine in modulating or optimizing the precision of prediction errors that may or may not be reward-related. (pp. 346-347)

Third, recent versions of the predictive processing theory argue that not only does the prediction error signal have various degrees of salience, but prior beliefs also have various degrees of salience.⁸⁷ The more salient a prior belief is (i.e. the more weighted, precise or certain it is), the less likely it will be updated with a different posterior belief. As Sterzer et al. (2018) put it, “an imprecise [i.e. less salient] prior will update more than a precise one will.” (p. 638) Crucially, the predictive processing theory argues that, like the salience of the prediction error signal, the salience of the prior belief is also modulated by certain neuromodulators such as dopamine (Corlett et al., 2016, p. 1147).

To recap, according to the predictive processing theory, the posterior belief is determined by the salience of the prior belief, and the salience of the prediction error signal. At the neurobiological level, the salience is modulated by certain neuromodulators such as dopamine (Marshall et al., 2016). The general point is that there is a salience-attribution mechanism in our brain and the malfunction of the salience-attribution mechanism is a cause of delusions (Howes et al., 2020).

⁸⁷ Early versions of the predictive processing theory were sometimes portrayed as a one-factor theory, and the only factor was the prediction error signal (e.g. Corlett et al., 2010b, p. 357). More recent versions of the predictive processing theory have included both the salience of the prediction error signal and the salience of priors. As Corlett et al. (2016, p. 1145) write in a review of the development of the predictive processing theory, “[w]hile previously we focused on the prediction error signal per se, an updated view takes into account its precision, as well as the precision of prior expectations.” This, however, raises the question of whether the updated version includes two separate factors (Sterzer et al., 2018, p. 637). I set aside this complication.

One possibility is, however, overlooked: that is, the salience-attribution mechanism may also directly misattribute salience to posterior (delusional) beliefs. If we call the predictive processing theory which only emphasises the salience of prediction error signals Salience Theory 1.0 (e.g. Corlett et al., 2007; Corlett et al., 2009; Corlett et al., 2010b); and call the predictive processing theory which adds the role of the salience of prior beliefs Salience Theory 2.0 (e.g. Corlett et al., 2016), we may call the present proposal which adds the role of the salience of posterior beliefs Salience Theory 3.0.

Of course, both Salience Theory 1.0 and Salience Theory 2.0 accept that the posterior delusional beliefs are salient to the subjects. However, for these two theories, the salience of posterior beliefs is determined by the salience of prediction error signals and the salience of prior beliefs. By contrast, Salience Theory 3.0 argues that the salience of posterior beliefs is not *completely* determined by the salience of prediction errors and the salience of priors, but can be directly influenced by the salience-attribution mechanism.⁸⁸

Why should we accept Salience Theory 3.0? One reason is that this possibility is already implied by the predictive processing theory's proposal of salience-attribution mechanism: there seems no reason the malfunctioning salience-attribution mechanism can only misattribute salience to prediction error signals and prior beliefs but not to posterior beliefs.

⁸⁸ Similar to Salience Theories 1.0 and 2.0, the two-factor theory appears to take it that the salience of a posterior belief is determined by antecedent factors such as the anomalous experience and the reasoning abnormality, and does not talk about the possibility that the salience of a posterior belief can have an independent source. For example, Aimola Davies et al. (2009) write: “[H]ow did the subject come to regard the false proposition as a salient and serious hypothesis, a credible candidate for belief? The first factor explains why the false proposition seemed a somewhat salient and credible hypothesis or why it was initially adopted as a belief.” (pp. 190-191) By contrast, I argue that the salience of the false proposition may be partly caused by something independent of the anomalous experience and the reasoning abnormality, a candidate for which may be the malfunctioning salience-attribution mechanism.

Another reason is that the salience of prediction error signals and prior beliefs falls short of explaining why delusional beliefs rather than other non-delusional beliefs are adopted by the subject. Consider the Capgras delusion. No matter how much the prediction error signal, i.e. the mismatch between the *expected* high-level autonomic activity and the *received* reduced autonomic activity, is over-weighted and how much the prior belief, i.e. that this woman is his wife, is underweighted, it still does not make sense for the subject to adopt the Capgras hypothesis that “This woman is an imposter” rather than other more plausible hypotheses, such as the brain-damage hypothesis and the fading-love hypothesis (for a similar critique of the predictive processing theory, see e.g. Parrott, 2021, p. 341).

How can the overlooked possibility that the content of the delusion is misattributed with salience help solve this problem with the predictive processing theory? From the perspective of the dual-force framework, the answer is that the delusional hypothesis is salient in the sense that it strikes the subject in such a phenomenally clear way that compels the subject’s assent.

I have linked the malfunctioning salience-attribution mechanism to the phenomenal clarity of the contents of delusions. Dopamine dysfunction has long been proposed to modulate the malfunctioning salience-attribution mechanism (Howes et al., 2020). If my analysis is on the right track, then it is likely that a certain kind of dopamine dysfunction may be a candidate for the neurobiological underpinning of the abnormal phenomenal clarity in the aetiology of delusions.⁸⁹

⁸⁹ I suggest that phenomenal clarity is a form of salience. But I do not claim that it is the only form of salience. As Ratcliffe and Broome (2018) argue, in the literature salience has been used as an umbrella concept that refers to a wide range of salient phenomena. Notably, Kapur (2003) argues that “[H]allucinations reflect a direct experience of the aberrant *salience* of internal representations.” (p. 13, emphasis added) From the perspective of the dual-force framework, the sense in which the content is saliently presented in a hallucination is quite different from the sense in which the content seems clear/salient to the subject to be true in their seeming

5. Conclusion

In this chapter, I have argued that the dual-force framework may advance our understanding of delusions with respect to appreciating the personal remarks on the resemblance between delusions and certain non-delusional beliefs, making some progress in addressing the evidence and specificity challenges, and shedding new light on the underpinnings of delusions. Can we now be confident to return to Jaspers and say that delusions are understandable? This is over-optimistic, however. There are some remaining questions that stand in our way to a complete understanding of delusions, to some of which the dual-force framework only offers partial answers.

First, the dual-force framework does not have a full answer to what I have called the content question: that is, why does a particular delusion have this rather than that content? It may have a partial answer: that is, the malfunctioning salience-attribution mechanism only attributes salience to this rather than that content such that only the content of the delusion strikes the person in a phenomenally clear way. But the dual-force framework has not told us why the malfunctioning salience-attribution mechanism only attributes salience to the content of the delusion, rather than to the content of a more plausible belief. More work is needed to answer the question of why the salience-attribution mechanism (mal)functions in this way.

Halligan and Marshall (1996) once remarked: “One would none the less hope that theories of normal belief-formation will eventually cast light on both the content of delusions and on the processes whereby the beliefs came to be held.” (p. 8) I wholeheartedly concur. The dual-force framework is developed on the basis of the formation

experience that gives rise to a delusion. One possible difference between them is that, unlike the clarity/salience of the contents of delusions, the salience of hallucinations may lack the same degree of phenomenal clarity and accordingly lack the same degree of the causal force that compels assent.

and maintenance of our normal beliefs (Chapter 3). Though it does not have a complete explanation of the contents of delusions, the progress it makes in explaining “the processes whereby the beliefs came to be held” helps keep our hopes up that there may be some overlooked resources in the formation and maintenance of our normal beliefs, which would eventually explain the contents of delusions. At the very least, I think the dual-force framework helps us remain hopeful that “delusions incomprehensible today may not be so tomorrow.” (Berrios, 1991, p. 10)

Second, the dual-force framework so far focuses on the interaction among clear experience, evidence, and (delusional) belief. It does not contain ready answers to the question of how memory, emotion, motivation, and action interact with (delusional) belief (see e.g. Bleuler, 1911/1950; Sass, 1994; Eilan, 2000; Davies et al., 2005; Davies, 2009). A tentative suggestion is that, like in some cases of akratic beliefs (Section 6 of Chapter 3), in some cases of delusions these factors may exert an impact on the phenomenal clarity of the subjects’ experiences.

Third, this thesis focuses on the monothematic, circumscribed, bizarre Capgras delusion, and I argue that clear experience is a contributory factor in this kind of delusions. More work is needed to explore to what extent clear experience is a necessary factor in a delusion that appears, in one way or another, different from the monothematic, circumscribed, bizarre Capgras delusion. Though what counts as a delusion itself can be a thorny issue, I acknowledge that clear experience might not be a necessary factor in some of what is conventionally taken as a delusion: such as a less severe form of paranoia that one’s neighbour is unfriendly or mean. This kind of paranoid belief is arguably understandable in the light of some not very anomalous experience and the jumping to conclusion bias (e.g. Ward & Garety, 2019).

Nevertheless, the dual-force framework can respect the convention and accommodate the diversity of delusions. From its perspective, delusions comprise a wide spectrum of phenomena. Towards one end, it is possible that the anomalous experience and/or reasoning abnormality suffice to outweigh the evidence against the delusion, where the subject may or may not have a clear experience; towards the other end, it is possible that the subject has no anomalous experience or reasoning abnormality and the only factor is their clear experience in which it seems clear to them that the content of the delusion is true; in between, I think, are the majority of delusional cases in which anomalous experiences, reasoning abnormalities and clear experiences work conjunctly to compel the subject's assent in the face of the counterevidence.

On top of these caveats concerning the explanatory power of the dual-force framework, we should also be cautious about taking the explanatory power of the dual-force framework as the guarantee that it is an infallible description of the fact about delusions. The explanatory power of a theory is normally the hallmark that the theory is close to the truth, but strictly speaking it is not a guarantee that the theory is true. For one thing, in the future there might be alternative theories which offer equally good or better explanations, and these alternative theories might or might not be compatible with the theory at issue. Consequently, these alternative theories would invite the question of which theory is true. For another, there is a possibility, albeit a remote one, that it is part of the nature of a certain matter of fact that it is inexplicable. Consequently, if a theory makes it explicable, then this theory might guide us away from the truth. That being said, before there is solid evidence that these remote possibilities raise serious concern about the values of the dual-force framework, I think we could remain confident that the dual-force framework helps us get a better grip on what is going on in delusions.

Overall, if the analysis in this thesis has been along the right lines, then it seems plausible that the dual-force framework can be taken as a new starting point for us to pursue knowledge about delusions and beliefs.

Appendix.

Are delusions beliefs?

1. Introduction

Are delusions really beliefs? When a person with the Capgras delusion asserts that his wife is an imposter (Capgras & Reboul-Lachaux, 1923; translated in Ellis et al., 1994; for reviews, see Edelstyn & Oyebode, 1999; Pandis et al., 2019), a person with the Cotard delusion says that she is dead (Cotard, 1882; for reviews, see Debruyne et al., 2009; Young & Leafhead, 1996), a person with the delusion of alien control asserts that the movement of her body is controlled by someone other than themselves (Frith, 1992; Frith & Done, 1989), do they really believe what they assert?⁹⁰

On the one hand, delusions strike us as so baffling and different from the paradigmatic beliefs, with which we are familiar in our daily lives, that it makes us wonder whether the subjects really believe them. On the other hand, delusions strike us as so baffling and different partly because they strike us as beliefs. If people with delusions were merely talking about some of their imagination, in a way like George R. R. Martin talking about Daenerys Targaryen's fire-breathing dragons and the Army of the Dead, then we would not be so much baffled. How do we resolve this tension? Are delusions really beliefs?

⁹⁰ This appendix is concerned with the question of whether delusions are beliefs. An answer to this question can be independent of the answer to the question of whether one's assertion is a delusion (Garety & Hemsley, 1994, chapter 1; Radden, 2011). When a patient makes an assertion, it is possible for some professionals to agree that the assertion is a delusion but disagree on whether delusions are beliefs; it is also possible for some professionals to agree that delusions are beliefs but disagree on whether the assertion is a delusion.

If you are uncertain of the meaning of the term *delusion*, as I was, look it up in the *Oxford English Dictionary*. It suggests that the primary meaning of delusion is a false belief. If you ask ordinary folks, they will probably tell you that delusions are beliefs (Rose et al., 2014).

If you want to know more about its meaning in clinical practice, open the *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)* (American Psychiatric Association, 2013). It reads: “Delusions are fixed beliefs that are not amenable to change in light of conflicting evidence.” (p. 87) If you ask mental health professionals, the majority will give their assent to it.^{91, 92} If you ask people with delusions, you will most likely find that they take their delusions as beliefs (Green et al., 2018).⁹³

If you skim through scientific journals, you will find that many scientific theories of delusions adopt the term *delusional beliefs* (e.g. Coltheart et al., 2011; Corlett et al., 2010b).

Now you may have become more certain that delusions are beliefs. Still, this quick survey may not have completely removed your worry. You may feel, and rightly so, that it is just conceptually baffling to take delusions as beliefs.

Naturally, with this conceptual bafflement you turn to philosophers. This time, however, you find no unanimous answer. Instead, you find yourself landing in the middle of a battlefield. On your right side is the doxasticist cohort who try to explain away your worry and argue that delusions are beliefs; on your left side is the anti-

⁹¹ There are notable exceptions. For example, Garety et al. (2020) hold that persecutory delusions “are frequently not fixed” beliefs, but “concerns, fears, worries, thoughts, [or] appraisals.” (p. 595)

⁹² This does not mean that it is always straightforward for mental health professionals to identify delusions with beliefs or the other way around. Related complications are evident in the Breivik case in Norway (Melle, 2013).

⁹³ Though people with delusions think their delusional beliefs are true, some of them are able to treat others’ delusional beliefs as false. For example, in Rokeach’s (1964/2011) controversial book, he reported that there were three patients with similar delusions: each patient believed that himself was Jesus Christ and the other two were not.

doxasticist cohort who not only share your worry but take it further and argue that delusions are something other than beliefs.

Is there a winning party yet? To answer that, we need to take a closer look at them: at their swords and guns and their armour and shields. As I will argue, it appears that the debate so far has encountered a deadlock. This deadlock has further fuelled the pessimistic views on the value of the inquiry into the doxastic nature of delusions and the value of folk psychology.

But be assured that it is not the aim of this appendix to torture you with more conceptual unsettlement and pessimistic views. In the end, with the help of the work developed in Chapters 3 and 4, I will try to convince you that folk psychology contains overlooked valuable resources that may help us get a better grip on what kind of beliefs delusions are. Before that, let us start by reviewing the puzzling features of delusions which initially fuel the worry about whether delusions are beliefs.

2. The puzzling features of delusions

In Chapter 1, we have discussed the flagrant ways delusions flout evidence and the specificity of delusions, which pose the two fundamental challenges to current theories of the aetiology of delusions. In this section, for the convenience of our discussion about the doxastic nature of delusions, I will briefly summarise some of the puzzling features of delusions we have seen in Chapter 1, and add several other puzzling features pertaining to our present discussion.

2.1. Insensitivity to evidence

The defining feature of delusions is their insensitivity to evidence (*DSM-5*). This means two things: delusions enjoy little supporting evidence and face significant counterevidence. Consider

the Cotard patient's assertion that she is dead: what could possibly be the supporting evidence? There is none. The counterevidence is, however, arguably overwhelming, such as the fact that the patient can walk and talk.

Other features are not as ubiquitously evident in all cases of delusions as the insensitivity to evidence, but may still strike us as something which is not typically associated with paradigmatic beliefs. Let us consider them in turn.

2.2. Cognitive inconsistency

Paradigmatic beliefs are, by and large, consistent with the subject's other beliefs. Delusions are, however, at odds with the patient's other beliefs. This phenomenon is most evident in monothematic delusions (Coltheart, 2013). Patients with monothematic delusions only have one delusion or several delusions closely related to one theme. They are "islands of delusion in a sea of apparent normality" (Davies & Egan, 2013, p. 690).

2.3. Incongruent intentions, affective states, and behaviours

Paradigmatic beliefs are often associated with congruent intentions and affective states. If someone believes that their spouse is missing, this belief is often associated with their intention to find their spouse and with their feeling distressed. Some patients with the Capgras delusion, however, show no intention to find their real spouses and do not feel distressed. Sometimes the patient even "had a friendly and warm relationship with" the imposter (Dietl et al., 2003, p. 462).

Some patients act on their delusions. Sometimes this can lead to violent tragedies. With the delusion that his stepfather is a robot, a patient decapitated his stepfather to look for "batteries and microfilm

in his head” (Blount, 1986, p. 207). But many patients do not act on their delusions. Many patients with the Capgras delusion do not go out to search for their real spouses or to hurt the imposters (Karakasi et al., 2019; Currell et al., 2019). This incongruence between what patients say and what they do has also been called behavioural asynchrony (Stone & Young, 1997, p. 334), double bookkeeping (Bleuler, 1911/1950; Gallagher, 2009; Sass, 1994; Sass, 2014), behavioural inertia (Frankish, 2009), and agential inertia (Wilkinson, 2013).

Why do these features of delusions raise doubts about whether delusions are beliefs? The reason can be spelt out in a number of different ways. According to dispositionalism, to say that a person possesses a belief is to say that the person possesses a certain cluster of dispositions related to that belief (Schwitzgebel, 2002). These puzzling features of delusions raise doubts about whether patients with delusions possess the characteristic dispositions (see e.g. Bayne & Pacherie, 2005; Tumulty, 2011). According to functionalism, what makes a mental state a belief is its functional roles. The puzzling features of delusions raise doubts about whether delusions have the characteristic functional roles of beliefs (see e.g. Currie, 2000; Wilkinson, 2013). According to interpretationism, what is important for us to ascribe a belief to a person is whether this belief can make the person’s actions or other mental states rational (Davidson, 1982/2004, 1985/2004; Dennett, 1987). The puzzling features of delusions raise doubts about whether delusions, if taken as beliefs, manifest too much irrationality (see e.g. Bortolotti, 2010).

The differences between these perspectives are important for many philosophical issues (Schwitzgebel, 2019). But I think, as I have implicitly assumed, it is clear that under the clothes of different philosophical perspectives they are talking about the same thing. The raw data are the same, so to speak (for similar views, see e.g.

Miyazono, 2019, p. 26; Clutton, 2018, pp. 208-210; Dub, 2017, pp. 34-35; Miyazono & Bortolotti, 2014, pp. 31-32; Porcher, 2019, pp. 114-115). Therefore, it seems OK if we undress them and look directly at the featural differences between delusions and paradigmatic beliefs, and only put the suitable clothes back when they pertain to a particular discussion.

The doubts about the doxastic nature of delusions, hence, can be summarised as follows:

- (1) Delusions are insensitive to evidence, inconsistent with the subjects' other beliefs, and associated with incongruent intentions, affective states, and behaviours.
- (2) Supposedly, paradigmatic beliefs are sensitive to evidence, consistent with the subjects' other beliefs, and associated with congruent intentions, affective states, and behaviours.
- (3) These featural differences raise doubts about whether delusions are beliefs.

3. The deadlock between doxasticists and anti-doxasticists

3.1. The arguments from the doxasticist cohort

So far, the most influential argument from the doxasticist cohort takes the strategy to revise (2). They argue that many of our everyday beliefs manifest delusion-like features, and the features of paradigmatic beliefs are improperly idealised. If the delusion-like features of these everyday beliefs do not prevent us from taking them as genuine beliefs, then, they argue, the features of delusions should not prevent us from taking delusions as beliefs. This strategy has been

spelt out in the parlance of interpretationism (Bortolotti, 2010), as well as dispositionalism and functionalism (Bayne & Pacherie, 2005; Miyazono & Bortolotti, 2014).⁹⁴

For interpretationists, the features of delusions manifest various kinds of irrationality. Bortolotti (2010) focuses on what she calls procedural irrationality, epistemic irrationality, and agential irrationality. Drawing on the studies of irrationality in empirical psychology, she argues that all the irrationalities manifested in delusions can be found in some everyday beliefs:

For each delusion, I'll give you a belief that matches the type if not the degree of irrationality of the delusion. ... [P]aranoid beliefs are badly supported by the evidence, and sometimes also as resistant to counterevidence as delusions of persecution. The delusion that I'm dead is often justified in a viciously circular way, not dissimilar from the way in which prejudiced beliefs against a racial group are justified. The delusion that someone of a higher status loves me, or the delusion that my partner is unfaithful can be defended by mentioning facts that are apparently irrelevant. This is not so different from the strenuous defence of superstitious beliefs which seem to get confirmation no matter what happens. (p. 259)^{95, 96}

⁹⁴ Philosopher's view may vary in different papers. For example, in her 2012 paper Bortolotti seems to emphasise the difference between the folk-psychological concept of belief and scientific concepts of belief (Bortolotti, 2012, p. 50), whereas in the 2015 paper co-authored by Miyazono and Bortolotti they seem to blur the difference and take these concepts as the same. In his 2008 paper Egan argues that delusions are intermediate attitudes between beliefs and imaginings, whereas in the 2013 paper co-authored by Davies and Egan they seem to assume that delusions are beliefs. Instead of trying to reconcile the apparent variation, I shall treat each paper as it stands.

⁹⁵ For the comparison between delusions and religious and political beliefs, see also Bentall (2018). For the comparison between delusions and philosophers' beliefs, see also Reimer (2010).

⁹⁶ Some doxasticists seem to hold that their argument stressing the delusion-like features of everyday beliefs only gives us a reason for rejecting that delusions are

3.2. The arguments from the anti-doxasticist cohort

Anti-doxasticists do not deny that many of our everyday beliefs have delusion-like features. In response, they may develop two distinct lines of argument. Anti-doxasticists may argue that neither delusions nor beliefs with delusion-like features are genuine beliefs. This argument is arguably too revisionary (for critique, see Bortolotti & Miyazono, 2015).

The other line of argument is less revisionary and more persuasive: anti-doxasticists can accept that everyday beliefs with delusion-like features are genuine beliefs but argue that the difference between the features of delusions and the delusion-like features of everyday beliefs are so drastic that delusions should not be taken as beliefs.

3.3. The deadlock

Now the debate between the doxasticist cohort and the anti-doxasticist cohort seems to focus on the question of to what extent the features of delusions differ from the features of beliefs. What is needed here is a theory that can both tell us what the necessary features of beliefs are, and tell us how to measure the featural differences between delusions and beliefs. Unfortunately, such a theory accepted by both sides does not exist. What both sides are offering are their intuitions. The intuitions are, however, in conflict with each other. The debate seems to have encountered a deadlock.

not beliefs, rather than gives us a reason for accepting that delusions are beliefs (e.g. Bortolotti, 2010, 2012; Miyazono & Bortolotti, 2014). Since the argument is, so far, the best argument from the doxastic cohort, it seems too weak to say that the argument does not give us a reason for accepting that delusions are beliefs. Does the weak claim mean that doxasticists have no positive philosophical reason for accepting that delusions are beliefs? It does not seem so. I shall take their argument as providing a positive reason too.

One attempt to break this deadlock is to shift our focus from the question of to what extent the features of delusions differ from the features of beliefs to the question of which kind of mental states can best capture the features of delusions.

For doxasticists, belief is a natural candidate since delusions have been taken as beliefs by default. What's more, the featural differences aside, there are many featural similarities between delusions and beliefs: like beliefs, delusions are sincerely asserted by patients; some delusions arguably show some sensitivity to evidence, albeit in a very limited sense; in some cases, the cognitive consistency is arguably preserved to a certain extent; some delusions are not associated with incongruent intentions, affective states, and behaviours.

These featural similarities, however, are not enough to convince the anti-doxasticist cohort. They argue that the features of delusions can be best captured by some non-belief mental states: imaginings (Currie, 2000; Currie & Jureidini, 2001; Currie & Ravenscroft, 2002); acceptances (Frankish, 2009; Dub, 2017); illusory experiences (Hohwy & Rajan, 2012; Hohwy, 2013); intermediate states between belief and imagination (Egan, 2008; Currie & Jones, 2006); intermediate states between belief and non-belief (Schwitzgebel, 2012; Tumulty, 2012, 2011).

For an illustration, let us consider Dub's (2017) reason for thinking of delusions as acceptances:

An acceptance that *p* is a disposition to draw conclusions from *p* in Type 2 reasoning; it is typically though not necessarily volitionally formed; the acceptance manifests itself in verbal behavior because we are often motivated to verbally express our acceptances in order to indirectly alter other mental states and to publicly affirm commitment or allegiance. These features explain why delusions are

circumscribed and exhibit double-bookkeeping. Patients are motivated to verbally express but not act on their acceptances, and they inchoately understand the distinction between their acceptances and beliefs, which explains the metacognitive awareness and phenomenological experience that seems to characterize double-bookkeeping. (p. 46)

Following the anti-doxasticists, we can add more candidates to this non-belief list. For example, a patient suffering from erotomania may have the delusion that a famous person is in love with her (Berrios & Kennedy, 2002). This erotomaniac delusion can be taken as an intermediate state between belief and desire: the patient partially believes that the famous person is in love with her and partially desires so.

Which of these proposed mental states, including beliefs, can best capture the features of delusions? To answer this question, we need a metatheory acting as an arbiter. Unfortunately, no such metatheory is accepted by both sides. Once again, the debate has encountered a deadlock.⁹⁷

This deadlock has increasingly been recognised by leading philosophers. For example, Bayne and Hattiangadi (2013) write:

[T]he reasons that motivate many people to classifying quasi-beliefs [i.e. beliefs with delusion-like features] as beliefs are also reasons for not classifying quasi-beliefs as instances of some other kind of standard mental state, such as imaginings. (p. 127)

Bortolotti and Miyazono (2015) write:

Revisionism [i.e. anti-doxasticists] and conservatism [i.e. doxasticists] can be both plausibly defended, and

⁹⁷ In fact, how easy it is to concoct another theory and how difficult it is to decide between them suggest that something is going wrong in the debate.

the budding literature on the topic shows that philosophers can make sense of many features of delusions independently of what strategy they opt for. (p. 639)

4. Several attempts to break the deadlock

Even though neither the doxasticist cohort nor the anti-doxasticist cohort has convinced each other when the features of delusions are concerned, there have been several other attempts to break the deadlock by focusing on something other than these features of delusions. Let us briefly consider them in turn.

4.1. Norms

Bayne (2010) argues that the intermediate-mental-state conception of delusions falls short of answering the question of what kind of norms govern these states. Regarding Egan's (2008) bimagination conception of delusions, according to which delusions are intermediate mental states between belief and imagination, Bayne (2010) argues that belief is governed by the norm of truth while imagination is not, but it is not clear in what sense bimagination is or is not governed by the norm of truth.⁹⁸

I think the unsettled relationship between bimagination and the norm of truth would be a problem for Egan's bimagination conception of delusions *if* he accepts that delusions are governed by the norms of truth. But I guess Egan does not have to accept that the concept of norms is applicable to delusions and bimaginations. He may reply that delusions and bimaginations precisely demonstrate the limitation

⁹⁸ For a similar critique of Schwitzgebel's (2012) conception of delusions as intermediate mental states between beliefs and non-beliefs, see Bayne and Hattiangadi (2013).

of the norm of truth. In fact, Egan could ask Bayne whether Bayne thinks delusions are governed by the norm of truth and how thinking of delusions as beliefs could help us settle the question.

Both the doxasticist cohort and the anti-doxasticist cohort can accept that belief is governed by certain doxastic norms while intermediate mental states are not. But neither has told us what kind of norms govern delusions, if there are any. Therefore, it is far from clear whether the consideration of norms can favour either side. The deadlock remains.

4.2. Cognitive feelings

Clutton (2018) provides what he calls “the cognitive phenomenological defence of doxasticism” (p. 198). Drawing on Kriegel (2015b) and others’ work on the experience of believing, Clutton (2018) argues that “there ... is a certain way it feels to be believing something” (p. 203). How could this help defend the doxastic view that the patient’s assertion that her spouse is an imposter is a belief? Clutton writes:

[O]ne nice move that the cognitive phenomenological view makes possible is to take ... [the patient’s] assertion, completely by itself, as strong evidence that she believes what she says. This is because the cognitive phenomenologist has a story about the means by which a person has acquaintance knowledge of their beliefs, namely, the phenomenology of the belief itself: the asserter is reporting on their experience [of the belief itself], and we have good reason to accept such reports. (p. 203)

Clutton is probably right that if patients’ experiences of delusions are embedded with a feeling of believing, then it is a good reason for taking delusions as beliefs. It is, however, far from clear

whether patients' experiences of delusions are embedded with cognitive feelings. Nor is it clear whether the feeling is the same as the feeling of believing, even if one accepts that the experiences of delusions are embedded with some cognitive feeling.

Dub (2017), from the anti-doxasticist cohort, agrees with Glutton that patients with delusions have some feeling. But, unlike Glutton, Dub argues that it is not a feeling embedded in the experiences of delusions but a feeling that gives rise to delusions. Dub writes:

When confronted with her spouse, the Capgras patient experiences a powerful feeling of unfamiliarity, typically due to a brain lesion. The feeling that the person before her is unfamiliar forces her to *accept* [rather than *believe*] that the person before her is unfamiliar. (p. 53, emphasis added)

Dub does not talk about whether patients' experiences of delusions are embedded with some cognitive feeling. But if we ask him, I guess he would be happy to accept that there is such a cognitive feeling and it is the feeling of accepting, rather than the feeling of believing.

Who is right? When patients assert their delusions, do they have the feeling of believing, or the feeling of accepting, or some other kind of feeling? Unlike ordinary experiences that are shared among ordinary folks such that it is *relatively* easy to reach an agreement on what experience a person is undergoing, in delusions patients' experiences are characteristically private and not shared with other people (see Hoerl, 2001; Radden, 2011). It is, therefore, unclear how non-deluded theorists can reach an agreement on what kind of

cognitive feeling is embedded in patients' private experiences.⁹⁹ The deadlock remains.¹⁰⁰

5. What next?

5.1. Two pessimistic views

The persistent deadlock has generated two pessimistic views: one is concerned with the value of the doxastic inquiry into whether delusions are beliefs; the other is concerned with the value of folk psychology, in particular the value of the folk-psychological concept of belief.

⁹⁹ Note that the issue is not about whether patients themselves take delusions as beliefs (Green et al., 2018), but about how an agreement can be reached about what experiences patients are undergoing.

¹⁰⁰ Another line of arguments proposed by some doxasticists is that scientific theories of delusions strongly support doxasticism. Why? They offer two reasons. One is that many scientific theories use the term *delusional beliefs* to refer to delusions. Call this the same-belief-term argument. The other reason, the doxasticists claim, is that according to scientific theories “delusions are ... produced by *exactly the same mechanisms* that produce paradigmatic beliefs.” (Miyazono & Bortolotti, 2014, p. 43, emphasis added; see also Clutton, 2018; Miyazono, 2019) Call this the same-mechanisms argument. Neither seems to be able to convince anti-doxasticists, however. First of all, there is a group of cognitive theories that are more sceptical about the doxastic nature of delusions (e.g. Garety et al., 2020; for a review, see e.g. Bell et al., 2006). The aforementioned papers by doxasticists surprisingly did not mention what Bell et al. (2006) call belief-negative theories of delusions. Regarding the same-belief-term argument, anti-doxasticists may argue that these scientists are not using the term *belief* in the same sense as the doxasticists. Rather, for scientists, the term *belief* is more like a placeholder (Murphy, 2012). Regarding the same-mechanisms argument, anti-doxasticists may argue that the difficulties of current theories of delusions in addressing the evidence challenge and the specificity challenge suggest that delusions may not be produced by “exactly the same mechanisms” that produce paradigmatic beliefs. Moreover, even with the help of the dual-force framework developed in this thesis, we still do not reach a complete understanding of delusions (Section 5 of Chapter 4). Therefore, for anti-doxasticists, this can be a reason to think that delusions are not produced by exactly the same mechanisms that produced paradigmatic beliefs. The deadlock remains.

5.1.1. The value of the doxastic inquiry

The debate on the question of whether delusions are beliefs can be dated back at least to Berrios' (1991, p. 12) suggestion that delusions are "empty speech acts". The literature has expanded rapidly since Currie's (2000) intriguing argument for taking delusions as imaginings. After decades of effort, however, there has been increasing doubt about the value of the debate.

From the doxasticist cohort, Bayne (2010) writes:

[W]ithout an account of the functional role of belief it is not clear whether this is really a debate about how best to understand delusions, as opposed to a debate about how to use the term 'belief.' (p. 332)

From the anti-doxasticist cohort, Hohwy (2013) writes:

It could be, then, that the way the deluded person deals with evidence convinces some theorists but not others to place delusions outside of the class of beliefs—but this would reflect people's respective theories of belief more than the nature of delusions. (p. 59)

Similar views are also held by philosophers who have not directly engaged in the debate. For example, Gerrans (2013) writes:

[O]nce the cognitive architecture of delusion formation is properly described the debate between doxastic and attitude [i.e. anti-doxastic] theorists loses its edge. (p. 83)

The pessimistic view is that the inquiry into whether delusions are beliefs seems to be a matter of terminology or personal preferences, and is inferior to the inquiry into the aetiology of delusions.

5.1.2. The value of folk psychology

The persistent deadlock has also fuelled the doubt on the value of folk psychology, in particular the value of the folk-psychological concept of belief (e.g. Porcher, 2018, 2019). Bayne (2010) writes:

I suspect that there may not be enough determinacy in our ordinary conception of belief for there to be a fact of the matter as to whether many belief-like states are really beliefs or not. (p. 332)

In his comment on Bortolotti's doxastic argument, Murphy (2012) writes:

[O]ur attributions of delusions may be guided by a sense that delusions are beliefs that we cannot explain in any folk psychological terms. (p. 19)

Similarly, Matthews (2013) writes:

I want to leave open the question whether they are genuine beliefs or maybe only belief-like, but I do intend to call attention to what seems to be a significant vagueness in our commonsense notion of belief, one that allows us to grade these states as more or less belief-like along a number of different dimensions. (p. 102)

According to this pessimistic view, it is the inherent limitation of our folk-psychological concept of belief that should be held accountable for the deadlock.

5.2. A diagnosis

Folk psychology, of course, has various limitations. But I think here it is unfair to lay the blame on folk psychology. It seems to me

that the real problem with the current debate is that both parties have mistakenly focused on comparing delusions with evidence-based beliefs.

While it is the defining *feature* of delusions that they are insensitive to evidence, it does not entail the *aetiological* claim that delusions are formed in response to little supporting evidence and maintained in the teeth of counterevidence. It could be a mistake to equate the feature with the aetiology. For example, it is characteristic of a martyr's belief of freedom that her belief of freedom is unwavering even when the price is as high as losing her life. But it could be a mistake to say that her unwavering belief is formed in response to, or in the teeth of, losing her life. In the literature on the doxastic nature of delusions, a conflation of the feature and aetiology of delusions seems to have infected both parties, though it is easier to discern in some writings than in others.

From the doxasticist cohort, Miyazono (2019) writes: “[I]t is widely assumed that doxastic *causal* roles include some degree of sensitivity to evidential inputs” (p. 23; emphasis added). This seems to be an aetiological claim. Right in the next paragraph, however, he writes: “[D]elusions are notoriously insensitive to counterevidence, and this insensitivity is often regarded as one of the defining *features* of delusions.” This seems to be a featural claim.

From the anti-doxastic cohort, Dub (2017) writes: “Delusions are *formed* in response to insufficient evidence, and they are *maintained* in spite of evidence to the contrary.” (p. 29; emphasis added) This seems to be an aetiological claim. Right in the next paragraph, however, he writes: “Unresponsiveness to evidence is one of the more well-known and commented-upon *features* of delusions. It is often presented as the defining *feature* that is constitutive of delusions.” This seems to be a featural claim.

I do not deny that the insensitivity to evidence and other puzzling features of delusions are important elements we should take

into account when we consider whether delusions are beliefs. But I think it is a mistake to think that evidence and how patients respond to evidence are the only important factors in the aetiology of delusions, and a mistake to think that whether delusions are beliefs is completely determined by the extent to which the features of delusions are like the features of evidence-based beliefs. As I will explain in the next section, the assumption that delusions are evidence-based beliefs may be what makes the features of delusion particularly puzzling, and more importantly many of the puzzling features can be explained away if we think of delusions as seeming-based beliefs.

6. Thinking of delusions as seeming-based beliefs

According to the dual-force framework developed in Chapters 3 and 4, apart from evidence, patients' clear experiences, in which it seems clear to them that the contents of their delusions are true, are also important factors in the aetiology of delusions. In comparison with thinking of delusions as evidence-based beliefs, we may think of delusions as partly seeming-based beliefs. Following this line of thought, whether delusions are beliefs is partly determined by to what extent delusions are like paradigmatic seeming-based beliefs, such as the belief that the sky is blue, the belief that we have five fingers on each hand, and the belief that 1 plus 1 is 2. Most often these beliefs are based on clear experiences in which it seems clear to us that they are true, rather than based on evidence.

Let us discuss in turn how thinking of delusions as seeming-based beliefs may help us explain some of the puzzling features of delusions.

First, it is particularly puzzling that delusions lack evidence when delusions are taken as evidence-based beliefs. But it is less puzzling if delusions are seeming-based beliefs. This is because in many cases of ordinary seeming-based beliefs, they lack evidence

too. For example, most often we believe that the sky is blue simply because it seems clear to us that the sky is blue, rather than because we have a lot of evidence that the sky is blue. Compare Sherlock Holmes' belief that Moriarty committed the crime with his belief that the sky is blue: it would be particularly puzzling if Holmes holds the former belief in the absence of evidence, but it would not be as puzzling if he holds the latter belief in the absence of evidence.

Second, it is puzzling that delusions face significant counterevidence, when delusions are taken as evidence-based beliefs. Admittedly, it still seems puzzling when we compare delusions with many ordinary seeming-based beliefs because most often our seeming-based beliefs are true, and they do not face counterevidence.

But, as Hobbes emphasised in his objection to Descartes, a clear experience that p is no guarantee that p is true, and it can be the reason why one obstinately believes something ("Thirteenth objection", in Descartes, 1984, p. 134). I have argued that akratic beliefs are better to be understood as a kind of seeming-based beliefs that face counterevidence (Section 6 of Chapter 3). If the presence of counterevidence does not prevent us from taking the Hobbesian obstinate beliefs in general or akratic beliefs in particular as genuine beliefs, then we have some reason to think that the presence of counterevidence does not suffice to prevent us from taking delusions as beliefs.

Third, regarding the cognitive inconsistency between delusions and patients' other beliefs, if we think that all beliefs are evidence-based, then it is particularly puzzling that patients are epistemically responsive to the evidence related to their non-delusional beliefs but somehow fail to be epistemically responsive to the evidence related to their delusions, and somehow fail to resolve the cognitive inconsistency between delusions and other non-delusional beliefs. If evidence and patients' epistemic ability to be responsive to evidence are the only factors at play, then these puzzles

suggest that patients with delusions suffer from a baffling form of epistemic irrationality.

By contrast, if delusions are seeming-based beliefs, then the compelling causal force of patients' seeming experiences, in which it seems clear that the contents of their delusions are true, can explain why they selectively fail to be epistemically responsive to the evidence related to their delusions, without appealing to some epistemic irrationality of the patients (Chapter 4). In fact, it may even be rational for patients to maintain the inconsistency between delusions and non-delusional beliefs so that they can minimise the damaging impact of seeming experiences.

Fourth, to say that delusions are evidence-based is, in a sense, to say that patients think that their delusions are somehow supported by evidence. If patients think that their delusions are somehow supported by evidence, then it is particularly puzzling why some patients do not have the intentions, affective states, or behaviours that are congruent with their delusions and congruent with what they take to be the supporting evidence.

By contrast, if delusions are seeming-based beliefs, then we can reject the assumption that patients think that their delusions are somehow supported by evidence. Instead, we can accept that some patients may recognise that their delusions enjoy little supporting evidence and face significant counterevidence. Following this line of thought, even though some patients' intentions, affective states, and behaviours are still not congruent with their seeming-based delusions, they may be congruent with their consideration of the evidence. Consider the patient with the Capgras delusion who does not intend to find his wife, does not feel distressed, and does not go out to search for his wife. Even though his intention, affective state, and behaviour are not congruent with his delusion, they seem to be congruent with his consideration of the little evidence related to his delusion. In the light of patients' consideration of the evidence related

to their delusions, it may even be rational for them to maintain the incongruence between their delusions and their intentions, affective states, and behaviours, so that they can maintain the congruence between their consideration of the evidence and their intentions, affective states, and behaviours.

If, as I have been trying to argue, comparing delusions to paradigmatic seeming-based beliefs that are familiar ones in folk psychology is a helpful way to understand why delusions are beliefs, then it seems that what has been taken as the limitation of folk psychology in explaining why delusions are beliefs is actually the limitation of thinking of delusions as evidence-based beliefs. Unlike the study of evidence-based beliefs, the study of seeming-based beliefs is relatively scarce in the literature, and a more comprehensive comparison between evidence-based beliefs and seeming-based beliefs can help us have a better understanding of the nature of (delusional) beliefs. All I am trying to illustrate here is that comparing delusions to paradigmatic seeming-based beliefs opens a promising line of inquiry into the doxastic nature of delusions.

To be clear, I do not claim that delusions are purely seeming-based beliefs. Nor do I claim that the puzzling features can never be reasons to doubt to what extent a delusion in a particular case is a belief. My point is rather modest. That is, thinking of delusions as seeming-based beliefs can help doxasticists explain why many delusions are beliefs. From the perspective of the dual-force framework, seeming experiences and evidence are two important factors that can have an impact on our beliefs; they may also act as distinct reasons that guide people's intentions, affective states, and behaviours. For a particular case of delusion, it is an empirical question to what extent the delusion is more like evidence-based beliefs or seeming-based beliefs.

In conclusion, if the analysis so far has been along the right lines, then it seems that the dual-force framework not only helps

advance our understanding of the aetiology of delusions, but, by offering plausible explanations of the puzzling features that otherwise fuel the doubt about whether delusions are beliefs, it also helps us get a better grip on why delusions are beliefs.

Bibliography

- Adams, R. A., Vincent, P., Benrimoh, D., Friston, K. J., & Parr, T. (2021). Everything is connected: Inference and attractors in delusions. *Schizophrenia Research*.
<https://doi.org/10.1016/j.schres.2021.07.032>
- Adler, J. E. (2002). Akratic believing? *Philosophical Studies*, 110(1), 1-27. <https://doi.org/10.1023/a:1019823330245>
- Aimola Davies, A. M., & Davies, M. (2009). Explaining pathologies of belief. In M. R. Broome & L. Bortolotti (Eds.), *Psychiatry as cognitive neuroscience* (pp. 284-324). Oxford University Press.
- Aimola Davies, A. M., Davies, M., Ogden, J. A., Smithson, M., & White, R. C. (2009). Cognitive and motivational factors in anosognosia. In T. Bayne & J. Fernandez (Eds.), *Delusion and self-deception: Affective and motivational influences on belief formation* (pp. 187-225). Psychology Press.
- Alston, W. P. (1988). The deontological conception of epistemic justification. *Philosophical Perspectives*, 2, 257-299.
<https://doi.org/10.2307/2214077>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (dsm-5)*. American Psychiatric Association.
- Antognazza, M. R. (2021). The distinction in kind between knowledge and belief. *Proceedings of the Aristotelian Society*, 120(3), 277-308. <https://doi.org/10.1093/arisoc/aoaa013>
- Ashinoff, B. K., Singletary, N. M., Baker, S. C., & Horga, G. (2021). Rethinking delusions: A selective review of delusion research through a computational lens. *Schizophrenia Research*.
<https://doi.org/10.1016/j.schres.2021.01.023>

- Asp, E., Khan, L., Jonason, A., Adkins-Hempel, M., Warner, K., Pardilla-Delgado, E., Payne, J., Mullineaux, P., Denburg, N., & Tranel, D. (2020). Second-guessing of spinoza: Psychophysiological and behavioral evidence that believing is default during proposition comprehension. *PsyArXiv*.
<https://doi.org/10.31234/osf.io/s3tfk>
- Austin, J. L. (1962). *Sense and sensibilia* (G. Warnock, Ed.). Oxford University Press.
- Aydede, M. (2019). Pain. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 ed.).
<https://plato.stanford.edu/archives/spr2019/entries/pain/>
- Ayers, M., & Antognazza, M. R. (2019). Knowledge and belief from plato to locke. In *Knowing and seeing* (pp. 3-33). Oxford University Press.
<https://doi.org/10.1093/oso/9780198833567.003.0001>
- Barton, J. J. S., Cherkasova, M. V., & Hefter, R. (2004). The covert priming effect of faces in prosopagnosia. *Neurology*, 63(11), 2062-2068.
<https://doi.org/10.1212/01.Wnl.0000145772.77040.24>
- Bate, S., & Cook, S. J. (2012). Covert recognition relies on affective valence in developmental prosopagnosia: Evidence from the skin conductance response. *Neuropsychology*, 26(5), 670-674. <https://doi.org/10.1037/a0029443>
- Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: A neuropsychological application of the guilty knowledge test. *Neuropsychologia*, 22(4), 457-469.
[https://doi.org/10.1016/0028-3932\(84\)90040-X](https://doi.org/10.1016/0028-3932(84)90040-X)
- Bayne, T. (2010). Delusions as doxastic states: Contexts, compartments, and commitments. *Philosophy, Psychiatry, & Psychology*, 17(4), 329-336.
<https://doi.org/10.1353/ppp.2010.0030>

- Bayne, T. (2017). Delusion and the norms of rationality. In T. J. Lane (Ed.), *Rationality: Constraints and contexts* (pp. 77-94). Academic Press. <https://doi.org/10.1016/b978-0-12-804600-5.00005-2>
- Bayne, T., & Hattiangadi, A. (2013). Belief and its bedfellows. In N. Nottelmann (Ed.), *New essays on belief: Constitution, content and structure* (pp. 124-144). Palgrave Macmillan. https://doi.org/10.1057/9781137026521_7
- Bayne, T., & Pacherie, E. (2004). Bottom-up or top-down: Campbell's rationalist account of monothematic delusions. *Philosophy, Psychiatry, & Psychology*, 11(1), 1-11. <https://doi.org/10.1353/ppp.2004.0033>
- Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusions. *Mind & Language*, 20(2), 163-188. <https://doi.org/10.1111/j.0268-1064.2005.00281.x>
- Bell, V., Halligan, P. W., & Ellis, H. D. (2006). Explaining delusions: A cognitive perspective. *Trends in Cognitive Sciences*, 10(5), 219-226. <https://doi.org/10.1016/j.tics.2006.03.004>
- Bell, V., Marshall, C., Kanji, Z., Wilkinson, S., Halligan, P., & Deeley, Q. (2017). Uncovering capgras delusion using a large-scale medical records database. *BJPsych Open*, 3(4), 179-185. <https://doi.org/10.1192/bjpo.bp.117.005041>
- Bell, V., Raihani, N., & Wilkinson, S. (2021). Derationalizing delusions. *Clin Psychol Sci*, 9(1), 24-37. <https://doi.org/10.1177/2167702620951553>
- Bengson, J. (2015). The intellectual given. *Mind*, 124(495), 707-760. <https://doi.org/10.1093/mind/fzv029>
- Bentall, R. P. (2018). Delusions and other beliefs. In L. Bortolotti (Ed.), *Delusions in context* (pp. 67-95). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-97202-2_3
- Bentall, R. P., Kaney, S., & Dewey, M. E. (1991). Paranoia and social reasoning: An attribution theory analysis. *British Journal of*

Clinical Psychology, 30 (Pt 1)(1), 13-23.

<https://doi.org/10.1111/j.2044-8260.1991.tb00915.x>

Berghofer, P. (2020a). Motivating and defending the phenomenological conception of perceptual justification. *Inquiry*, 1-18.

<https://doi.org/10.1080/0020174x.2020.1712232>

Berghofer, P. (2020b). Towards a phenomenological conception of experiential justification. *Synthese*, 197(1), 155-183.

<https://doi.org/10.1007/s11229-018-1744-5>

Bermudez, J. L. (2001). Normativity and rationality in delusional psychiatric disorders. *Mind & Language*, 16(5), 493-457.

<https://doi.org/10.1111/1468-0017.00179>

Berrios, G. E. (1991). Delusions as “wrong beliefs”: A conceptual history. *British Journal of Psychiatry*, 159(S14), 6-13.

<https://doi.org/10.1192/s0007125000296414>

Berrios, G. E., & Kennedy, N. (2002). Erotomania: A conceptual history. *History of Psychiatry*, 13(52), 381-400.

<https://doi.org/10.1177/0957154X0201305202>

Bleuler, P. E. (1911/1950). *Dementia praecox or the group of schizophrenias* (J. Zinkin, Trans.). International Universities Press.

Blount, G. (1986). Dangerousness of patients with capgras syndrome. *Nebraska Medical Journal*, 71(6), 207.

Bobes, M. A., Lopera, F., Comas, L. D., Galan, L., Carbonell, F., Bringas, M. L., & Valdes-Sosa, M. (2004). Brain potentials reflect residual face processing in a case of prosopagnosia. *Cognitive Neuropsychology*, 21(7), 691-718.

<https://doi.org/10.1080/02643290342000258>

Bongiorno, F. (2019). Is the capgras delusion an endorsement of experience? *Mind & Language*, 35(3), 293-312.

<https://doi.org/10.1111/mila.12239>

- Bongiorno, F. (2021). Spinozan doxasticism about delusions. *Pacific Philosophical Quarterly*. <https://doi.org/10.1111/papq.12373>
- BonJour, L. (1985). *The structure of empirical knowledge*. Harvard University Press.
- Borgoni, C., & Luthra, Y. (2017). Epistemic akrasia and the fallibility of critical reasoning. *Philosophical Studies*, 174(4), 877-886. <https://doi.org/10.1007/s11098-016-0711-6>
- Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford University Press.
- Bortolotti, L. (2012). In defence of modest doxasticism about delusions. *Neuroethics*, 5(1), 39-53. <https://doi.org/10.1007/s12152-011-9122-8>
- Bortolotti, L. (2015). The epistemic innocence of motivated delusions. *Consciousness and Cognition*, 33, 490-499. <https://doi.org/10.1016/j.concog.2014.10.005>
- Bortolotti, L. (2016). Epistemic benefits of elaborated and systematized delusions in schizophrenia. *British Journal for the Philosophy of Science*, 67(3), 879-900. <https://doi.org/10.1093/bjps/axv024>
- Bortolotti, L. (2020). *The epistemic innocence of irrational beliefs*. Oxford University Press.
- Bortolotti, L., & Broome, M. R. (2008). Delusional beliefs and reason giving. *Philosophical Psychology*, 21(6), 821-841. <https://doi.org/10.1080/09515080802516212>
- Bortolotti, L., & Miyazono, K. (2015). Recent work on the nature and development of delusions. *Philosophy Compass*, 10(9), 636-645. <https://doi.org/10.1111/phc3.12249>
- Bortolotti, L., Sullivan-Bissett, E., & Gunn, R. (2016). What makes a belief delusional? In I. McCarthy, K. Sellevold, & O. Smith (Eds.), *Cognitive confusions: Dreams, delusions and illusions in early modern culture* (pp. 37-51). Modern Humanities Research Association.

- Braun, C. M. J., & Suffren, S. (2011). A general neuropsychological model of delusion. *Cognitive Neuropsychiatry*, 16(1), 1-39.
<https://doi.org/10.1080/13546800903442314>
- Breen, N., Caine, D., & Coltheart, M. (2000a). Models of face recognition and delusional misidentification: A critical review. *Cognitive Neuropsychology*, 17(1), 55-71.
<https://doi.org/10.1080/026432900380481>
- Breen, N., Caine, D., Coltheart, M., Hendy, J., & Roberts, C. (2000b). Towards an understanding of delusions of misidentification: Four case studies. *Mind & Language*, 15(1), 74-110.
<https://doi.org/10.1111/1468-0017.00124>
- Breen, N., Coltheart, M., & Caine, D. (2001). A two-way window on face recognition. *Trends in Cognitive Sciences*, 5(6), 234-235.
[https://doi.org/10.1016/s1364-6613\(00\)01659-4](https://doi.org/10.1016/s1364-6613(00)01659-4)
- Brewer, B. (1999). *Perception and reason*. Clarendon Press.
- Brighetti, G., Bonifacci, P., Borlimi, R., & Ottaviani, C. (2007). "Far from the heart far from the eye": Evidence from the capgras delusion. *Cognitive Neuropsychiatry*, 12(3), 189-197.
<https://doi.org/10.1080/13546800600892183>
- Brogaard, B. (2013). Phenomenal seemings and sensible dogmatism. In C. Tucker (Ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism* (pp. 270-289). Oxford University Press.
- Brogaard, B., & Gatzia, D. E. (2017). The real epistemic significance of perceptual learning. *Inquiry*, 61(5-6), 543-558.
<https://doi.org/10.1080/0020174x.2017.1368172>
- Bronstein, M. V., Pennycook, G., Joormann, J., Corlett, P. R., & Cannon, T. D. (2019). Dual-process theory, conflict processing, and delusional belief. *Clinical Psychology Review*, 72, Article 101748. <https://doi.org/10.1016/j.cpr.2019.101748>
- Broome, M. R. (2004). The rationality of psychosis and understanding the deluded. *Philosophy, Psychiatry, &*

Psychology, 11(1), 35-41.

<https://doi.org/10.1353/ppp.2004.0035>

Bruce, V., & Young, A. (1986). Understanding face recognition.

British Journal of Psychology, 77 (Pt 3)(3), 305-327.

<https://doi.org/10.1111/j.2044-8295.1986.tb02199.x>

Bruce, V., & Young, A. (2012). *Face perception*. Psychology Press.

Bruyer, R. (1991). Covert face recognition in prosopagnosia: A review. *Brain and Cognition*, 15(2), 223-235.

[https://doi.org/10.1016/0278-2626\(91\)90027-6](https://doi.org/10.1016/0278-2626(91)90027-6)

Campbell, J. (1999). Schizophrenia, the space of reasons, and thinking as a motor process. *Monist*, 82(4), 609-625.

<https://doi.org/10.5840/monist199982426>

Campbell, J. (2001). Rationality, meaning, and the analysis of delusion. *Philosophy, Psychiatry, & Psychology*, 8(2), 89-100.

<https://doi.org/10.1353/ppp.2001.0004>

Campbell, J. (2002). *Reference and consciousness*. Oxford University Press.

Campbell, J. (2007). An interventionist approach to causation in psychology. In A. Gopnik & L. J. Schulz (Eds.), *Causal learning: Psychology, philosophy and computation* (pp. 58-66). Oxford University Press.

Campbell, J. (2008). Causation in psychiatry. In K. S. Kendler & J. Parnas (Eds.), *Philosophical issues in psychiatry: Explanation, phenomenology, and nosology* (pp. 196-216). John Hopkins University Press.

Campbell, J. (2013). Causation and mechanisms in psychiatry. In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, & T. Thornton (Eds.), *Oxford handbook of philosophy and psychiatry* (pp. 935-949). Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780199579563.013.0055>

- Campbell, J. (2014). The role of sensory experience in propositional knowledge. In *Berkeley's puzzle: What does experience teach us?* (pp. 76-99). Oxford University Press.
- Campbell, J. (2020). *Causation in psychology*. Harvard University Press.
- Capgras, J., & Reboul-Lachaux, J. (1923). L'illusion des "sosies" dans un délire systématisé chronique. *Bulletin de la Société Clinique de Médecine Mentale*, 2, 6–16.
- Carriero, J. (2013). X-epistemology past and present. *Proceedings of the Aristotelian Society*, 113(2pt2), 175-200.
<https://doi.org/10.1111/j.1467-9264.2013.00351.x>
- Cassam, Q. (2007). Ways of knowing. *Proceedings of the Aristotelian Society*, 107(1), 339-358. <https://doi.org/10.1111/j.1467-9264.2007.00225.x>
- Cassam, Q. (2014). Experientialism. In *Berkeley's puzzle: What does experience teach us?* (pp. 119-136). Oxford University Press.
- Cassam, Q. (2019). *Conspiracy theories*. Polity Press.
- Cassam, Q. (2020). *Philosophy for the real world: Knowledge, ignorance and bullshit [lecture 3, class handout]*. University of Warwick, Coventry, UK.
- Cermolacce, M., Sass, L., & Parnas, J. (2010). What is bizarre in bizarre delusions? A critical review. *Schizophrenia Bulletin*, 36(4), 667-679. <https://doi.org/10.1093/schbul/sbq001>
- Chasid, A., & Weksler, A. (2020). Belief-like imaginings and perceptual (non-)assertoricity. *Philosophical Psychology*, 33(5), 731-751.
<https://doi.org/10.1080/09515089.2020.1751104>
- Christodoulou, G. N. (1986). The origin of the concept of 'doubles'. In G. N. Christodoulou (Ed.), *The delusional misidentification syndromes* (pp. 1-8). Karger.

- Chudnoff, E. (2011a). The nature of intuitive justification. *Philosophical Studies*, 153(2), 313-333.
<https://doi.org/10.1007/s11098-010-9495-2>
- Chudnoff, E. (2011b). What intuitions are like. *Philosophy and Phenomenological Research*, 82(3), 625-654.
<https://doi.org/10.1111/j.1933-1592.2010.00463.x>
- Chudnoff, E. (2012). Presentational phenomenology. In S. Miguens & G. Preyer (Eds.), *Consciousness and subjectivity* (pp. 51-72). Ontos Verlag.
- Chudnoff, E. (2013). *Intuition*. Oxford University Press.
- Church, J. (2013). *Possibilities of perception*. Oxford University Press.
- Cipriani, G., Danti, S., Vedovello, M., Nuti, A., & Lucetti, C. (2014). Understanding delusion in dementia: A review. *Geriatrics & Gerontology International*, 14(1), 32-39.
<https://doi.org/10.1111/ggi.12105>
- Clamor, A., & Krkovic, K. (2018). Paranoid delusions as an adaptive response to social evaluative stress? *Zeitschrift für Psychologie*, 226(3), 191-196. <https://doi.org/10.1027/2151-2604/a000332>
- Clutton, P. (2018). A new defence of doxasticism about delusions: The cognitive phenomenological defence. *Mind & Language*, 33(2), 198-217. <https://doi.org/10.1111/mila.12164>
- Colombo, M., Elkin, L., & Hartmann, S. (2021). Being realist about bayes, and the predictive processing theory of mind. *The British Journal for the Philosophy of Science*, 72(1), 185-220.
<https://doi.org/10.1093/bjps/axy059>
- Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences*, 3(3), 115-120. [https://doi.org/10.1016/S1364-6613\(99\)01289-9](https://doi.org/10.1016/S1364-6613(99)01289-9)
- Coltheart, M. (2005). Conscious experience and delusional belief. *Philosophy, Psychiatry, & Psychology*, 12(2), 153-157.

- Coltheart, M. (2007). The 33rd Bartlett lecture: Cognitive neuropsychiatry and delusional belief. *Quarterly Journal of Experimental Psychology*, 60(8), 1041-1062.
<https://doi.org/10.1080/17470210701338071>
- Coltheart, M. (2010). The neuropsychology of delusions. *Annals of the New York Academy of Sciences*, 1191(1), 16-26.
<https://doi.org/10.1111/j.1749-6632.2010.05496.x>
- Coltheart, M. (2013). On the distinction between monothematic and polythematic delusions. *Mind & Language*, 28(1), 103-112. <https://doi.org/10.1111/mila.12011>
- Coltheart, M. (2017). Confabulation and conversation. *Cortex*, 87, 62-68. <https://doi.org/10.1016/j.cortex.2016.08.002>
- Coltheart, M., Cox, R., Sowman, P., Morgan, H., Barnier, A., Langdon, R., Connaughton, E., Teichmann, L., Williams, N., & Polito, V. (2018). Belief, delusion, hypnosis, and the right dorsolateral prefrontal cortex: A transcranial magnetic stimulation study. *Cortex*, 101, 234-248.
<https://doi.org/10.1016/j.cortex.2018.01.001>
- Coltheart, M., & Davies, M. (2021a). Failure of hypothesis evaluation as a factor in delusional belief. *Cognitive Neuropsychiatry*, 26(4), 213-230.
<https://doi.org/10.1080/13546805.2021.1914016>
- Coltheart, M., & Davies, M. (2021b). How unexpected observations lead to new beliefs: A Peircean pathway. *Consciousness and Cognition*, 87. <https://doi.org/10.1016/j.concog.2020.103037>
- Coltheart, M., Langdon, R., & McKay, R. (2007). Schizophrenia and monothematic delusions. *Schizophrenia Bulletin*, 33(3), 642-647. <https://doi.org/10.1093/schbul/sbm017>
- Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, 62(1), 271-298.
<https://doi.org/10.1146/annurev.psych.121208.131622>

- Coltheart, M., Menzies, P., & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive Neuropsychiatry*, 15(1), 261-287. <https://doi.org/10.1080/13546800903439120>
- Combs, D. R., Adams, S. D., Michael, C. O., Penn, D. L., Basso, M. R., & Gouvier, W. D. (2006). The conviction of delusional beliefs scale: Reliability and validity. *Schizophrenia Research*, 86(1-3), 80-88. <https://doi.org/10.1016/j.schres.2006.06.023>
- Connors, M. H. (2015). Hypnosis and belief: A review of hypnotic delusions. *Consciousness and Cognition*, 36, 27-43. <https://doi.org/10.1016/j.concog.2015.05.015>
- Connors, M. H., & Halligan, P. W. (2020). Delusions and theories of belief. *Consciousness and Cognition*, 81, Article 102935. <https://doi.org/10.1016/j.concog.2020.102935>
- Corlett, P. R. (2018). Delusions and prediction error. In L. Bortolotti (Ed.), *Delusions in context* (pp. 35-66). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-97202-2_2
- Corlett, P. R. (2019). Factor one, familiarity and frontal cortex: A challenge to the two-factor theory of delusions. *Cognitive Neuropsychiatry*, 24(3), 165-177. <https://doi.org/10.1080/13546805.2019.1606706>
- Corlett, P. R. (2021). Turning away from the social turn. *Imperfect Cognitions*. <http://imperfectcognitions.blogspot.com/2021/06/social-approaches-to-delusions-5.html>
- Corlett, P. R., D'Souza, D. C., & Krystal, J. H. (2010a). Capgras syndrome induced by ketamine in a healthy subject. *Biological Psychiatry*, 68(1), e1-2. <https://doi.org/10.1016/j.biopsych.2010.02.015>
- Corlett, P. R., & Fletcher, P. (2021). Modelling delusions as temporally-evolving beliefs. *Cognitive Neuropsychiatry*, 26(4), 231-241. <https://doi.org/10.1080/13546805.2021.1938984>

- Corlett, P. R., & Fletcher, P. C. (2014). Computational psychiatry: A rosetta stone linking the brain to mental illness. *Lancet Psychiatry*, 1(5), 399-402. [https://doi.org/10.1016/S2215-0366\(14\)70298-6](https://doi.org/10.1016/S2215-0366(14)70298-6)
- Corlett, P. R., & Fletcher, P. C. (2015). Delusions and prediction error: Clarifying the roles of behavioural and brain responses. *Cognitive Neuropsychiatry*, 20(2), 95-105. <https://doi.org/10.1080/13546805.2014.990625>
- Corlett, P. R., Honey, G. D., & Fletcher, P. C. (2016). Prediction error, ketamine and psychosis: An updated model. *Journal of Psychopharmacology*, 30(11), 1145-1155. <https://doi.org/10.1177/0269881116650087>
- Corlett, P. R., Krystal, J. H., Taylor, J. R., & Fletcher, P. C. (2009). Why do delusions persist? *Frontiers in Human Neuroscience*, 3, Article 12. <https://doi.org/10.3389/neuro.09.012.2009>
- Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R., Shanks, D. R., Robbins, T. W., Bullmore, E. T., Dickinson, A., & Fletcher, P. C. (2007). Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain*, 130(Pt 9), 2387-2400. <https://doi.org/10.1093/brain/awm173>
- Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010b). Toward a neurobiology of delusions. *Progress in Neurobiology*, 92(3), 345-369. <https://doi.org/10.1016/j.pneurobio.2010.06.007>
- Cotard, J. (1882). Du délire des négations. *Archives de Neurologie*, 4, 152-170, 282-195.
- Cottingham, J. (Ed.). (1992). *The cambridge companion to descartes*. Cambridge University Press.
- Cottingham, J. (1994). *Reason, will, and sensation: Studies in descartes's metaphysics*. Clarendon Press.
- Cottingham, J. (Ed.). (1998). *Descartes*. Oxford University Press.

- Crane, T., & French, C. (2021). The problem of perception. In N. Z. Edwards (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 ed.).
<https://plato.stanford.edu/archives/fall2021/entries/perception-problem/>
- Currell, E. A., Werbeloff, N., Hayes, J. F., & Bell, V. (2019). Cognitive neuropsychiatric analysis of an additional large capgras delusion case series. *Cognitive Neuropsychiatry*, 24(2), 123-134. <https://doi.org/10.1080/13546805.2019.1584098>
- Currie, G. (2000). Imagination, delusion and hallucinations. *Mind & Language*, 15(1), 168-183. <https://doi.org/10.1111/1468-0017.00128>
- Currie, G., & Jones, N. (2006). McGinn on delusion and imagination. *Philosophical Books*, 47(4), 306-313.
https://doi.org/10.1111/j.1468-0149.2006.410_2.x
- Currie, G., & Jureidini, J. (2001). Delusion, rationality, empathy: Commentary on martin davies et al. *Philosophy, Psychiatry, & Psychology*, 8(2), 159-162.
<https://doi.org/10.1353/ppp.2001.0006>
- Currie, G., & Ravenscroft, I. (2002). *Recreative minds: Imagination in philosophy and psychology*. Oxford University Press.
- Cutting, J. (2003). Descriptive psychopathology. In S. R. Hirsch & D. R. Weinberger (Eds.), *Schizophrenia* (2nd ed., pp. 15-24). Blackwell Science.
- Dalgalarondo, P., Fujisawa, G., & Banzato, C. E. (2002). Capgras syndrome and blindness: Against the prosopagnosia hypothesis. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, 47(4), 387-388.
<https://doi.org/10.1177/070674370204700421>
- Darby, R., & Prasad, S. (2016). Lesion-related delusional misidentification syndromes: A comprehensive review of reported cases. *Journal of Neuropsychiatry and Clinical*

Neurosciences, 28(3), 217-222.

<https://doi.org/10.1176/appi.neuropsych.15100376>

Darby, R. R., & Fox, M. D. (2017). Reply: Capgras syndrome: Neuroanatomical assessment of brain mri findings in an adolescent patient. *Brain*, 140(7), Article e44.

<https://doi.org/10.1093/brain/awx125>

Darby, R. R., Laganieri, S., Pascual-Leone, A., Prasad, S., & Fox, M. D. (2017). Finding the imposter: Brain connectivity of lesions causing delusional misidentifications. *Brain*, 140(2), 497-507.

<https://doi.org/10.1093/brain/aww288>

David, A. S. (1990). Insight and psychosis. *British Journal of Psychiatry*, 156(06), 798-808.

<https://doi.org/10.1192/bjp.156.6.798>

David, A. S. (1999). On the impossibility of defining delusions.

Philosophy, Psychiatry, & Psychology, 6(1), 17-20.

David, A. S., & Halligan, P. W. (1996). Editorial. *Cognitive Neuropsychiatry*, 1(1), 1-4.

<https://doi.org/10.1080/135468096396659>

David, A. S., & Prince, M. (2005). Psychosis following head injury: A critical review. *Journal of Neurology, Neurosurgery and Psychiatry*, 76(suppl 1), i53-i60.

<https://doi.org/10.1136/jnnp.2004.060475>

Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy*, 60(23), 685-700. <https://doi.org/10.2307/2023177>

Davidson, D. (1969). How is weakness of the will possible? In J. Feinberg (Ed.), *Moral concepts*. Oxford University Press.

Davidson, D. (1982/2004). Paradoxes of irrationality. In *Problems of rationality* (pp. 169-187). Oxford University Press.

(Philosophical Essays on Freud, edited by R. Wollheim and J. Hopkins (Cambridge University Press, 1982), 289-305)

- Davidson, D. (1985/2004). Incoherence and irrationality. In *Problems of rationality* (pp. 189-198). Oxford University Press.
(*Dialectica*, 39 (1985), 345–54.)
- Davidson, D. (1986). A coherence theory of truth and knowledge. In E. LePore (Ed.), *Truth and interpretation: Perspectives on the philosophy of donald davidson* (pp. 307–319). Basil Blackwell.
- Davies, M. (2000a). Interaction without reduction: The relationship between personal and sub-personal levels of description. *Mind and Society*, 1(2), 87-105.
<https://doi.org/10.1007/bf02512315>
- Davies, M. (2000b). Persons and their underpinnings. *Philosophical Explorations*, 3(1), 43-62.
<https://doi.org/10.1080/13869790008520980>
- Davies, M. (2009). Delusion and motivationally biased belief: Self-deception in the two-factor framework. In T. Bayne & J. Fernandez (Eds.), *Delusion and self-deception: Affective and motivational influences on belief formation* (pp. 71-86). Psychology Press.
- Davies, M., & Coltheart, M. (2000). Introduction: Pathologies of belief. *Mind & Language*, 15(1), 1-46.
<https://doi.org/10.1111/1468-0017.00122>
- Davies, M., Coltheart, M., Langdon, R., & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry, & Psychology*, 8(2), 133-158.
<https://doi.org/10.1353/ppp.2001.0007>
- Davies, M., Davies, A. A., & Coltheart, M. (2005). Anosognosia and the two-factor theory of delusions. *Mind & Language*, 20(2), 209-236. <https://doi.org/10.1111/j.0268-1064.2005.00283.x>
- Davies, M., & Egan, A. (2013). Delusion: Cognitive approaches—bayesian inference and compartmentalisation. In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, & T. Thornton (Eds.), *The oxford handbook of*

philosophy and psychiatry (pp. 689-727). Oxford University Press.

Davies, M., McGill, C. L., & Aimola Davies, A. M. (in press).

Anosognosia for motor impairments as a delusion: Anomalies of experience and belief evaluation. In A. L. Mishara, P. R. Corlett, P. C. Fletcher, A. Kranjec, & M. A. Schwartz (Eds.), *Phenomenological neuropsychiatry: How patient experience bridges clinic with clinical neuroscience*. Springer.

de Pauw, K. W. (1994). Psychodynamic approaches to the capgras delusion: A critical historical review. *Psychopathology*, 27(3-5), 154-160. <https://doi.org/10.1159/000284864>

Debruyne, H., Portzky, M., Van den Eynde, F., & Audenaert, K. (2009). Cotard's syndrome: A review. *Current Psychiatry Reports*, 11(3), 197-202. <https://doi.org/10.1007/s11920-009-0031-z>

Dellsén, F. (2017). The heuristic conception of inference to the best explanation. *Philosophical Studies*, 175(7), 1745-1766. <https://doi.org/10.1007/s11098-017-0933-2>

Dennett, D. C. (1986). *Content and consciousness* (2nd ed.). Routledge. (1969, First Edition)

Dennett, D. C. (1987). *The intentional stance*. MIT Press.

Descartes, R. (1984). *The philosophical writings of descartes, vol. 2* (J. Cottingham, R. Stoothoff, & D. Murdoch, Trans.). Cambridge University Press.

<https://doi.org/10.1017/cbo9780511818998>

Descartes, R. (1991). *The philosophical writings of descartes: The correspondence, vol 3* (J. Cottingham, R. Stoothoff, D. Murdoch, & A. Kenny, Trans.). Cambridge University Press.

Dewey, J. (1929). *Experience and nature* (2nd ed.). Open Court Publishing Company.

Diamond, B. J., Valentine, T., Mayes, A. R., & Sandel, M. E. (1994). Evidence of covert recognition in a prosopagnosic patient.

Cortex, 30(3), 377-393. [https://doi.org/10.1016/S0010-9452\(13\)80336-X](https://doi.org/10.1016/S0010-9452(13)80336-X)

Diaz-Cutraro, L., Garcia-Mieres, H., Lopez-Carrilero, R., Ferrer, M., Verdaguer-Rodriguez, M., Barrigon, M. L., Barajas, A., Grasa, E., Pousa, E., Lorente, E., Ruiz-Delgado, I., Gonzalez-Higueras, F., Cid, J., Palma-Sevillano, C., Moritz, S., Spanish Metacognition, G., & Ochoa, S. (2021). Jumping to conclusions is differently associated with specific subtypes of delusional experiences: An exploratory study in first-episode psychosis. *Schizophrenia Research*, 228, 357-359.

<https://doi.org/10.1016/j.schres.2020.12.037>

Dietl, T., Herr, A., Brunner, H., & Friess, E. (2003). Capgras syndrome--out of sight, out of mind? *Acta Psychiatrica Scandinavica*, 108(6), 460-462; discussion 462-463.

<https://doi.org/10.1046/j.0001-690x.2003.00135.x>

Dub, R. (2017). Delusions, acceptances, and cognitive feelings. *Philosophy and Phenomenological Research*, 94(1), 27-60.

<https://doi.org/10.1111/phpr.12220>

Dubourg, E., & Baumard, N. (2021). Why imaginary worlds? *Behavioral and Brain Sciences*, 1-52.

<https://doi.org/10.1017/S0140525X21000923>

Dudley, R., Taylor, P., Wickham, S., & Hutton, P. (2016). Psychosis, delusions and the "jumping to conclusions" reasoning bias: A systematic review and meta-analysis. *Schizophrenia Bulletin*, 42(3), 652-665. <https://doi.org/10.1093/schbul/sbv150>

Dupont, J. (1988). Ferenczi's "madness". *Contemporary Psychoanalysis*, 24(2), 250-261.

<https://doi.org/10.1080/00107530.1988.10746240>

Eames, S. M. (1964). Primary experience in the philosophy of john dewey. *Monist*, 48(3), 407-418.

<https://doi.org/10.5840/monist196448325>

- Edelstyn, N. M., & Oyebode, F. (1999). A review of the phenomenology and cognitive neuropsychological origins of the capgras syndrome. *International Journal of Geriatric Psychiatry*, 14(1), 48-59. [https://doi.org/10.1002/\(SICI\)1099-1166\(199901\)14:1<48::AID-GPS891>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1099-1166(199901)14:1<48::AID-GPS891>3.0.CO;2-0)
- Egan, A. (2008). Imagination, delusion, and self-deception. In T. Bayne & J. Fernández (Eds.), *Delusions, self-deception: Affective and motivational influences on belief formation* (pp. 263–280). Psychology Press.
- Eilan, N. (2000). On understanding schizophrenia. In D. Zahavi (Ed.), *Exploring the self: Philosophical and psychopathological perspectives on self-experience* (pp. 97-113). John Benjamins Publishing.
- Eilan, N. (2001). Meaning, truth, and the self: Commentary on campbell and parnas and sass. *Philosophy, Psychiatry, & Psychology*, 8(2), 121-132. <https://doi.org/10.1353/ppp.2001.0008>
- Elliott, B., Joyce, E., & Shorvon, S. (2009a). Delusions, illusions and hallucinations in epilepsy: 1. Elementary phenomena. *Epilepsy Research*, 85(2-3), 162-171. <https://doi.org/10.1016/j.eplepsyres.2009.03.018>
- Elliott, B., Joyce, E., & Shorvon, S. (2009b). Delusions, illusions and hallucinations in epilepsy: 2. Complex phenomena and psychosis. *Epilepsy Research*, 85(2-3), 172-186. <https://doi.org/10.1016/j.eplepsyres.2009.03.017>
- Ellis, H. D. (1989). Past and recent studies of prosopagnosia. In J. R. Crawford & D. M. Parker (Eds.), *Developments in clinical and experimental neuropsychology* (pp. 151-166). Springer. https://doi.org/10.1007/978-1-4757-9996-5_13
- Ellis, H. D., & de Pauw, K. W. (1994). The cognitive neuropsychiatric origins of the capgras delusion. In A. S. David & J. C. Cutting

- (Eds.), *The neuropsychology of schizophrenia* (pp. 317-335). Psychology Press.
- Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: A window on face recognition. *Trends in Cognitive Sciences*, 5(4), 149-156. [https://doi.org/10.1016/S1364-6613\(00\)01620-X](https://doi.org/10.1016/S1364-6613(00)01620-X)
- Ellis, H. D., Lewis, M. B., Moselhy, H. F., & Young, A. W. (2000). Automatic without autonomic responses to familiar faces: Differential components of covert face recognition in a case of capgras delusion. *Cognitive Neuropsychiatry*, 5(4), 255-269. <https://doi.org/10.1080/13546800050199711>
- Ellis, H. D., Whitley, J., & Luaute, J. P. (1994). Delusional misidentification: The three original papers on the capgras, fregoli and intermetamorphosis delusions. *History of Psychiatry*, 5(17 Pt 1), 117-146. <https://doi.org/10.1177/0957154X9400501708>
- Ellis, H. D., & Young, A. W. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry*, 157(2), 239-248. <https://doi.org/10.1192/bjp.157.2.239>
- Ellis, H. D., Young, A. W., Quayle, A. H., & De Pauw, K. W. (1997). Reduced autonomic responses to faces in capgras delusion. *Proceedings: Biological Sciences*, 264(1384), 1085-1092. <https://doi.org/10.1098/rspb.1997.0150>
- Enoch, D., Puri, B. K., & Ball, H. (2020). Othello syndrome. In *Uncommon psychiatric syndromes* (pp. 51-73). Routledge.
- Factor, S. A., Scullin, M. K., Sollinger, A. B., Land, J. O., Wood-Siverio, C., Zanders, L., Freeman, A., Bliwise, D. L., McDonald, W. M., & Goldstein, F. C. (2014). Cognitive correlates of hallucinations and delusions in parkinson's disease. *Journal of the Neurological Sciences*, 347(1-2), 316-321. <https://doi.org/10.1016/j.jns.2014.10.033>
- Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of*

Experimental Psychology: General, 144(5), 993-1002.

<https://doi.org/10.1037/xge0000098>

Ferguson, M. R., Yu, C. K., Poliakov, A. V., Friedman, S. D., & McClellan, J. M. (2017). Capgras syndrome: Neuroanatomical assessment of brain mri findings in an adolescent patient. *Brain*, 140(7), Article e43.

<https://doi.org/10.1093/brain/awx121>

Feyaerts, J., Henriksen, M. G., Vanheule, S., Myin-Germeys, I., & Sass, L. A. (2021a). Delusions beyond beliefs: A critical overview of diagnostic, aetiological, and therapeutic schizophrenia research from a clinical-phenomenological perspective. *Lancet Psychiatry*, 8(3), 237-249.

[https://doi.org/10.1016/S2215-0366\(20\)30460-0](https://doi.org/10.1016/S2215-0366(20)30460-0)

Feyaerts, J., Kusters, W., Van Duppen, Z., Vanheule, S., Myin-Germeys, I., & Sass, L. (2021b). Uncovering the realities of delusional experience in schizophrenia: A qualitative phenomenological study in belgium. *Lancet Psychiatry*, 8(9), 784-796. [https://doi.org/10.1016/S2215-0366\(21\)00196-6](https://doi.org/10.1016/S2215-0366(21)00196-6)

Fine, C., Craigie, J., & Gold, I. (2005). Damned if you do; damned if you don't: The impasse in cognitive accounts of the capgras delusion. *Philosophy, Psychiatry, & Psychology*, 12(2), 143-151.

Fine, C., Gardner, M., Craigie, J., & Gold, I. (2007). Hopping, skipping or jumping to conclusions? Clarifying the role of the jtc bias in delusions. *Cognitive Neuropsychiatry*, 12(1), 46-77.

<https://doi.org/10.1080/13546800600750597>

Fineberg, S. K., & Corlett, P. R. (2016). The doxastic shear pin: Delusions as errors of learning and memory. *Cognitive Neuropsychiatry*, 21(1), 73-89.

<https://doi.org/10.1080/13546805.2015.1136206>

Flores, C. (2021). Delusional evidence-responsiveness. *Synthese*.

<https://doi.org/10.1007/s11229-021-03070-2>

- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.
- Franceschi, P. (2010). A logical defence of maher's model of polythematic delusions. *Journal of Philosophical Research*, 35(2). (English translation and postprint of a paper originally published in French in *Philosophiques*, 2008, vol. 35-2, pp. 451-475)
- Frankish, K. (2009). Delusions: A two-level framework. In M. R. Broome & L. Bortolotti (Eds.), *Psychiatry as cognitive neuroscience: Philosophical perspectives* (pp. 269-284). Oxford University Press.
- Freedman, K. L. (2017). Akrotic believing, psychological trauma, and somatic representations. *Philosophy, Psychiatry, & Psychology*, 24(4), 337-346.
<https://doi.org/10.1353/ppp.2017.0049>
- Freeman, D., & Garety, P. (2014). Advances in understanding and treating persecutory delusions: A review. *Social Psychiatry and Psychiatric Epidemiology*, 49(8), 1179-1189.
<https://doi.org/10.1007/s00127-014-0928-7>
- Friston, K. (2012). The history of the future of the bayesian brain. *NeuroImage*, 62(2), 1230-1233.
<https://doi.org/10.1016/j.neuroimage.2011.10.004>
- Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Psychology Press.
- Frith, C. D., & Done, D. J. (1989). Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychological Medicine*, 19(2), 359-363.
<https://doi.org/10.1017/s003329170001240x>
- Fulford, K. W. M. (1989). *Moral theory and medical practice*. Cambridge University Press.
- Gallagher, S. (2009). Delusional realities. In M. R. Broome & L. Bortolotti (Eds.), *Psychiatry as cognitive neuroscience*:

Philosophical perspectives (pp. 245-268). Oxford University Press.

Gallagher, S. (2018). Phenomenology and the cognitive sciences. In G. Stanghellini, M. Broome, A. Raballo, A. V. Fernandez, P. Fusar-Poli, & R. Rosfort (Eds.), *The oxford handbook of phenomenological psychopathology* (pp. 262-272). Oxford University Press.

<https://doi.org/10.1093/oxfordhb/9780198803157.013.24>

Gangemi, A., & Cardella, V. (2014). Reasoning in schizophrenia. In N. Galbraith (Ed.), *Aberrant beliefs and reasoning* (pp. 99-113). Psychology Press.

Garety, P. A., & Freeman, D. (2013). The past and future of delusions research: From the inexplicable to the treatable. *British Journal of Psychiatry*, 203(5), 327-333.

<https://doi.org/10.1192/bjp.bp.113.126953>

Garety, P. A., & Hemsley, D. R. (1994). *Delusions: Investigations into the psychology of delusional reasoning*. Oxford University Press.

Garety, P. A., Hemsley, D. R., & Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients. Biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease*, 179(4), 194-201.

<https://doi.org/10.1097/00005053-199104000-00003>

Garety, P. A., Ward, T., & Rus-Calafell, M. (2020). Beyond belief—new approaches to the treatment of paranoia. In J. C. Badcock & G. Paulik (Eds.), *A clinical introduction to psychosis* (pp. 591-613). Academic Press.

<https://doi.org/10.1016/B978-0-12-815012-2.00025-0>

Genin, K., & Huber, F. (2021). Formal representations of belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 ed.).

<https://plato.stanford.edu/archives/fall2021/entries/formal-belief/>

- Gerrans, P. (2002). A one-stage explanation of the cotard delusion. *Philosophy, Psychiatry, & Psychology*, 9(1), 47-53. <https://doi.org/10.1353/ppp.2003.0007>
- Gerrans, P. (2013). Delusional attitudes and default thinking. *Mind & Language*, 28(1), 83-102. <https://doi.org/10.1111/mila.12010>
- Gerrans, P. (2014). *The measure of madness: Philosophy of mind, cognitive neuroscience, and delusional thought*. MIT Press.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46(2), 107-119. <https://doi.org/10.1037/0003-066x.46.2.107>
- Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, 59(4), 601-613. <https://doi.org/10.1037/0022-3514.59.4.601>
- Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, 65(2), 221-233. <https://doi.org/10.1037/0022-3514.65.2.221>
- Ginet, C. (2001). Deciding to believe. In M. Steup (Ed.), *Knowledge, truth and duty* (pp. 63-76). Oxford University Press.
- Gödel, K. (1984). What is cantor's continuum problem? In H. Putnam & P. Benacerraf (Eds.), *Philosophy of mathematics: Selected readings* (2 ed., pp. 470-485). Cambridge University Press. <https://doi.org/10.1017/CBO9781139171519.025>
- Gold, I., & Hohwy, J. (2000). Rationality and schizophrenic delusion. *Mind & Language*, 15(1), 146-167. <https://doi.org/10.1111/1468-0017.00127>
- Gold, J., & Gold, I. (2012). The "truman show" delusion: Psychosis in the global village. *Cognitive Neuropsychiatry*, 17(6), 455-472. <https://doi.org/10.1080/13546805.2012.666113>

- Gold, J., & Gold, I. (2015). *Suspicious minds: How culture shapes madness*. Free Press.
- Greco, D. (2014). A puzzle about epistemic akrasia. *Philosophical Studies*, 167(2), 201-219. <https://doi.org/10.1007/s11098-012-0085-3>
- Green, H., Hauser, L., & Troyakov, V. (2018). Are delusions beliefs? A qualitative examination of the doxastic features of delusions. *Psychosis*, 10(4), 319-328. <https://doi.org/10.1080/17522439.2018.1528298>
- Gunn, R., & Bortolotti, L. (2018). Can delusions play a protective role? *Phenomenology and the Cognitive Sciences*, 17(4), 813-833. <https://doi.org/10.1007/s11097-017-9555-6>
- Gunn, R., & Larkin, M. (2019). Delusion formation as an inevitable consequence of a radical alteration in lived experience. *Psychosis*, 12(2), 151-161. <https://doi.org/10.1080/17522439.2019.1690562>
- Halligan, P. W., & Marshall, J. C. (Eds.). (1996). *Method in madness: Case studies in cognitive neuropsychiatry*. Psychology Press.
- Halligan, P. W., Marshall, J. C., & Wade, D. T. (1993). Three arms: A case study of supernumerary phantom limb after right hemisphere stroke. *Journal of Neurology, Neurosurgery and Psychiatry*, 56(2), 159-166. <https://doi.org/10.1136/jnnp.56.2.159>
- Halpern, J. Y. (2017). *Reasoning about uncertainty* (2nd ed.). MIT Press.
- Harrison, E. (2019). Seemings, truth-makers, and epistemic justification. *Synthese*, 198(6), 5689-5708. <https://doi.org/10.1007/s11229-019-02428-x>
- Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological Science*, 16(7), 566-571. <https://doi.org/10.1111/j.0956-7976.2005.01576.x>

- Heck, R. G. (2000). Nonconceptual content and the "space of reasons". *The Philosophical Review*, 109(4), 483-523.
<https://doi.org/10.2307/2693622>
- Heil, J. (1984). Doxastic incontinence. *Mind*, XCIII(369), 56-70.
<https://doi.org/10.1093/mind/XCIII.369.56>
- Henriksen, M. G., & Parnas, J. (2018). Delusional mood. In G. Stanghellini, M. Broome, A. Raballo, A. V. Fernandez, P. Fusar-Poli, & R. Rosfort (Eds.), *The oxford handbook of phenomenological psychopathology* (pp. 743-752). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780198803157.013.72>
- Hermanowicz, N. (2002). A blind man with parkinson's disease, visual hallucinations, and capgras syndrome. *Journal of Neuropsychiatry and Clinical Neurosciences*, 14(4), 462-463.
<https://doi.org/10.1176/jnp.14.4.462>
- Hirstein, W., & Ramachandran, V. S. (1997). Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons. *Proceedings: Biological Sciences*, 264(1380), 437-444.
<https://doi.org/10.1098/rspb.1997.0062>
- Hoerl, C. (2001). Introduction: Understanding, explaining, and intersubjectivity in schizophrenia. *Philosophy, Psychiatry, & Psychology*, 8(2), 83-88.
<https://doi.org/10.1353/ppp.2001.0010>
- Hoerl, C. (2019). Understanding and explaining in psychiatry. In G. Stanghellini, M. Broome, A. V. Fernandez, P. Fusar-Poli, A. Raballo, & R. Rosfort (Eds.), *The oxford handbook of phenomenological psychopathology* (pp. 407-413). Oxford University Press.
- Hohwy, J. (2004). Top-down and bottom-up in delusion formation. *Philosophy, Psychiatry, & Psychology*, 11(1), 65-70.
<https://doi.org/10.1353/ppp.2004.0043>

- Hohwy, J. (2013). Delusions, illusions and inference under uncertainty. *Mind & Language*, 28(1), 57-71.
<https://doi.org/10.1111/mila.12008>
- Hohwy, J., & Rajan, V. (2012). Delusions as forensically disturbing perceptual inferences. *Neuroethics*, 5(1), 5-11.
<https://doi.org/10.1007/s12152-011-9124-6>
- Hohwy, J., & Rosenberg, R. (2005). Unusual experiences, reality testing and delusions of alien control. *Mind & Language*, 20(2), 141-162. <https://doi.org/10.1111/j.0268-1064.2005.00280.x>
- Hookway, C. (2001). Epistemic akrasia and epistemic virtue. In A. Fairweather & L. Zagzebski (Eds.), *Virtue epistemology: Essays on epistemic virtue and responsibility* (pp. 178-199). Oxford University Press.
- Horne, Z., Muradoglu, M., & Cimpian, A. (2019). Explanation as a cognitive process. *Trends in Cognitive Sciences*, 23(3), 187-199. <https://doi.org/10.1016/j.tics.2018.12.004>
- Howes, O. D., Hird, E. J., Adams, R. A., Corlett, P. R., & McGuire, P. (2020). Aberrant salience, information processing, and dopaminergic signaling in people at clinical high risk for psychosis. *Biological Psychiatry*, 88(4), 304-314.
<https://doi.org/10.1016/j.biopsych.2020.03.012>
- Huemer, M. (2001). *Skepticism and the veil of perception*. Rowman & Littlefield Publishers.
- Huemer, M. (2005). *Ethical intuitionism*. Palgrave Macmillan.
<https://doi.org/10.1057/9780230597051>
- Huemer, M. (2019). Phenomenal conservatism. In *Internet Encyclopedia of Philosophy*. <https://www.iep.utm.edu/phen-con/>
- Hume, D. (1739/2007). *A treatise of human nature* (D. F. Norton & M. J. Norton, Eds. Vol. 1: Texts). Clarendon Press.

- Humpston, C. S., Adams, R. A., Benrimoh, D., Broome, M. R., Corlett, P. R., Gerrans, P., Horga, G., Parr, T., Pienkos, E., Powers, A. R., 3rd, Raballo, A., Rosen, C., & Linden, D. E. J. (2019). From computation to the first-person: Auditory-verbal hallucinations and delusions of thought interference in schizophrenia-spectrum psychoses. *Schizophrenia Bulletin*, 45(Supplement_1), S56-S66.
<https://doi.org/10.1093/schbul/sby073>
- Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *The Quarterly Journal of Experimental Psychology Section A*, 40(4), 801-812. <https://doi.org/10.1080/14640748808402300>
- Husserl, E. (1960). *Cartesian meditations: An introduction to phenomenology* (D. Cairns, Trans.). Martinus Nijhoff.
- Hyman, J. (2015). *Action, knowledge, and will*. Oxford University Press.
- James, W. (1890/1950). *The principles of psychology*. Henry Holt & Co.
- Janet, P. (1929). *De l'angoisse à l'extase*. Felix Alcan.
- Jardri, R., Duverne, S., Litvinova, A. S., & Denève, S. (2017). Experimental evidence for circular inference in schizophrenia. *Nature Communications*, 8(1), Article 14218.
<https://doi.org/10.1038/ncomms14218>
- Jaspers, K. (1913/1997). *General psychopathology* (J. Hoenig & M. W. Hamilton, Trans.). Johns Hopkins University Press.
- Jolley, S., & Garety, P. A. (2004). Insight and delusions: A cognitive psychological approach. In X. F. Amador & A. S. David (Eds.), *Insight and psychosis: Awareness of illness in schizophrenia and related disorders* (2nd ed., pp. 89-100). Oxford University Press.
- Jones, E., & Watson, J. P. (2018). Delusion, the overvalued idea and religious beliefs: A comparative analysis of their

- characteristics. *British Journal of Psychiatry*, 170(04), 381-386. <https://doi.org/10.1192/bjp.170.4.381>
- Jones, R. D., & Tranel, D. (2001). Severe developmental prosopagnosia in a child with superior intellect. *Journal of Clinical and Experimental Neuropsychology*, 23(3), 265-273. <https://doi.org/10.1076/jcen.23.3.265.1183>
- Joseph, Z., Victor, K., & Rimona, D. (2011). "Ego-dystonic" delusions as a predictor of dangerous behavior. *Psychiatric Quarterly*, 82(2), 113-120. <https://doi.org/10.1007/s11126-010-9150-2>
- Josephs, K. A. (2007). Capgras syndrome and its relationship to neurodegenerative disease. *Archives of Neurology*, 64(12), 1762-1766. <https://doi.org/10.1001/archneur.64.12.1762>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.
- Kapur, N., Young, A., Bateman, D., & Kennedy, P. (1989). Focal retrograde amnesia: A long term clinical and neuropsychological follow-up. *Cortex*, 25(3), 387-402. [https://doi.org/10.1016/S0010-9452\(89\)80053-X](https://doi.org/10.1016/S0010-9452(89)80053-X)
- Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, 160(1), 13-23. <https://doi.org/10.1176/appi.ajp.160.1.13>
- Karakasi, M. V., Markopoulou, M., Alexandri, M., Douzenis, A., & Pavlidis, P. (2019). In fear of the most loved ones. A comprehensive review on capgras misidentification phenomenon and case report involving attempted murder under capgras syndrome in a relapse of a schizophrenia

- spectrum disorder. *Journal of Forensic and Legal Medicine*, 66, 8-24. <https://doi.org/10.1016/j.jflm.2019.05.019>
- Kearl, T. (2019). Epistemic akrasia and higher-order beliefs. *Philosophical Studies*, 177(9), 2501-2515. <https://doi.org/10.1007/s11098-019-01323-y>
- Keeley, B. L. (1999). Of conspiracy theories. *The Journal of Philosophy*, 96(3), 109-126. <https://doi.org/10.2307/2564659>
- Kim, H. R., Han, J. Y., Park, Y. H., Kim, B. J., Yang, W., & Kim, S. (2017). Supernumerary phantom limb in a patient with basal ganglia hemorrhage - a case report and review of the literature. *BMC Neurology*, 17(1), Article 180. <https://doi.org/10.1186/s12883-017-0962-7>
- Kind, A. M. Y. (2017). Imaginative vividness. *Journal of the American Philosophical Association*, 3(1), 32-50. <https://doi.org/10.1017/apa.2017.10>
- Klee, R. (2004). Delusional content and the public nature of meaning: Reply to the other contributors. *Philosophy, Psychiatry, & Psychology*, 11(1), 95-99. <https://doi.org/10.1353/ppp.2004.0045>
- Koksvik, O. (2011). *Intuition* [Ph.D. Thesis, The Australian National University].
- Koritar, E., & Steiner, W. (1988). Capgras' syndrome: A synthesis of various viewpoints. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, 33(1), 62-66. <https://doi.org/10.1177/070674378803300115>
- Kraepelin, E. (1889). *Psychiatrie*. Abel.
- Kriegel, U. (2015a). Perception and imagination. In S. Miguens, G. Preyer, & C. Bravo Morando (Eds.), *Prereflective consciousness: Sartre and contemporary philosophy of mind* (pp. 245-276). Routledge.
- Kriegel, U. (2015b). *The varieties of consciousness*. Oxford University Press.

- Krstić, V. (2019). Transparent delusion. *Review of Philosophy and Psychology*, 11(1), 183-201. <https://doi.org/10.1007/s13164-019-00457-6>
- Lancellotta, E., & Bortolotti, L. (2019). Are clinical delusions adaptive? *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(5), Article e1502. <https://doi.org/10.1002/wcs.1502>
- Langdon, R., & Bayne, T. (2010). Delusion and confabulation: Mistakes of perceiving, remembering and believing. *Cognitive Neuropsychiatry*, 15(1), 319-345. <https://doi.org/10.1080/13546800903000229>
- Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind & Language*, 15(1), 184-218. <https://doi.org/10.1111/1468-0017.00129>
- Lewis, M., & Ellis, H. (2001). A two-way window on face recognition: Reply to Breen et al. *Trends in Cognitive Sciences*, 5(6), 235-235. [https://doi.org/10.1016/s1364-6613\(00\)01666-1](https://doi.org/10.1016/s1364-6613(00)01666-1)
- Lewis, M. B., Sherwood, S., Moselhy, H., & Ellis, H. D. (2001). Autonomic responses to familiar faces without autonomic responses to familiar voices: Evidence for voice-specific capgras delusion. *Cognitive Neuropsychiatry*, 6(3), 217-228. <https://doi.org/10.1080/13546800143000041>
- Liu, J., Wang, M., Shi, X., Feng, L., Li, L., Thacker, J. M., Tian, J., Shi, D., & Lee, K. (2014). Neural correlates of covert face processing: Fmri evidence from a prosopagnosic patient. *Cerebral Cortex*, 24(8), 2081-2092. <https://doi.org/10.1093/cercor/bht059>
- Liu, M., & Klein, C. (2020). Pain and spatial inclusion: Evidence from mandarin. *Analysis*, 80(2), 262-272. <https://doi.org/10.1093/analys/anz032>
- Livet, A., Navarri, X., Potvin, S., & Conrod, P. (2020). Cognitive biases in individuals with psychotic-like experiences: A systematic review and a meta-analysis. *Schizophrenia*

Research, 222, 10-22.

<https://doi.org/10.1016/j.schres.2020.06.016>

- Locke, J. (1690/1975). *An essay concerning human understanding* (P. H. Nidditch, Ed.). Oxford University Press.
<https://doi.org/10.1093/actrade/9780198243861.book.1> (First published 1690)
- Longworth, G. (2020). John langshaw austin. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020 ed.).
<https://plato.stanford.edu/archives/fall2020/entries/austin-jl/>
- López-Silva, P. (2018). Mapping the psychotic mind: A review on the subjective structure of thought insertion. *Psychiatric Quarterly*, 89(4), 957-968. <https://doi.org/10.1007/s11126-018-9593-4>
- Mackie, J., Ebmeier, K. P., & O'Carroll, R. E. (1994). An mri, spect and neuropsychological study of a patient presenting with capgras syndrome. *Behavioural Neurology*, 7, 211-215.
<https://doi.org/10.3233/BEN-1994-73-417>
- Maher, B. A. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology*, 30(1), 98-113.
- Maher, B. A. (1988). Anomalous experience and delusional thinking: The logic of explanations. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs* (pp. 15-33). John Wiley and Sons.
- Maher, B. A. (1992). Delusions: Contemporary etiological hypotheses. *Psychiatric Annals*, 22(5), 260-268.
<https://doi.org/10.3928/0048-5713-19920501-11>
- Maher, B. A. (1999). Anomalous experience in everyday life: Its significance for psychopathology. *Monist*, 82(4), 547-570.
<https://doi.org/10.5840/monist199982428>
- Maher, B. A. (2001). Delusions. In P. B. Sutker & H. E. Adams (Eds.), *Comprehensive handbook of psychopathology* (pp. 309-339). Springer.

- Maher, B. A. (2006). The relationship between delusions and hallucinations. *Current Psychiatry Reports*, 8(3), 179-183.
<https://doi.org/10.1007/s11920-006-0021-3>
- Mandelbaum, E. (2014). Thinking is believing. *Inquiry*, 57(1), 55-96.
<https://doi.org/10.1080/0020174X.2014.858417>
- Mandik, P. (2007). Shit happens. *Episteme*, 4(2), 205-218.
<https://doi.org/10.3366/epi.2007.4.2.205>
- Marshall, L., Mathys, C., Ruge, D., de Berker, A. O., Dayan, P., Stephan, K. E., & Bestmann, S. (2016). Pharmacological fingerprints of contextual uncertainty. *PLoS Biology*, 14(11), Article e1002575.
<https://doi.org/10.1371/journal.pbio.1002575>
- Matthen, M. (2005). *Seeing, doing, and knowing: A philosophical theory of sense perception*. Oxford University Press.
- Matthews, R. J. (2013). Belief and belief's penumbra. In N. Nottelmann (Ed.), *New essays on belief: Constitution, content and structure* (pp. 100-123). Palgrave Macmillan.
https://doi.org/10.1057/9781137026521_6
- McDowell, J. (1994). *Mind and world*. Harvard University Press.
- McGrath, M., & Frank, D. (2020). Propositions. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2020 ed.).
<https://plato.stanford.edu/archives/win2020/entries/propositions/>
- McKay, R. (2012). Delusional inference. *Mind & Language*, 27(3), 330-355. <https://doi.org/10.1111/j.1468-0017.2012.01447.x>
- McKay, R. (2019). Measles, magic and misidentifications: A defence of the two-factor theory of delusions. *Cognitive Neuropsychiatry*, 24(3), 183-190.
<https://doi.org/10.1080/13546805.2019.1607273>
- McKay, R., & Dennett, D. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, 32(6), 493-510; discussion 510-461. <https://doi.org/10.1017/S0140525X09990975>

- McKay, R., Langdon, R., & Coltheart, M. (2007). Jumping to delusions? Paranoia, probabilistic reasoning, and need for closure. *Cognitive Neuropsychiatry*, 12(4), 362-376.
<https://doi.org/10.1080/13546800701203769>
- McKay, R. T., & Ross, R. M. (2020). Religion and delusion. *Current Opinion in Psychology*, 40, 160-166.
<https://doi.org/10.1016/j.copsyc.2020.10.002>
- McLean, B. F., Mattiske, J. K., & Balzan, R. P. (2017). Association of the jumping to conclusions and evidence integration biases with delusions in psychosis: A detailed meta-analysis. *Schizophrenia Bulletin*, 43(2), 344-354.
<https://doi.org/10.1093/schbul/sbw056>
- Mele, A. R. (1986). Incontinent believing. *The Philosophical Quarterly*, 36(143), 212-222. <https://doi.org/10.2307/2219769>
- Mele, A. R. (2001). *Self-deception unmasked*. Princeton University Press.
- Melle, I. (2013). The breivik case and what psychiatrists can learn from it. *World Psychiatry*, 12(1), 16-21.
<https://doi.org/10.1002/wps.20002>
- Millar, A. (2011). How visual perception yields reasons for belief. *Philosophical Issues*, 21, 332-351.
<https://doi.org/10.1111/j.1533-6077.2011.00207.x>
- Mishara, A. L., & Corlett, P. (2010). Are delusions biologically adaptive? Salvaging the doxastic shear pin. *Behavioral and Brain Sciences*, 32(6), 530-531.
<https://doi.org/10.1017/s0140525x09991464>
- Miyazono, K. (2019). *Delusions and beliefs: A philosophical inquiry*. Routledge.
- Miyazono, K., & Bortolotti, L. (2014). The causal role argument against doxasticism about delusions. *AVANT. The Journal of the Philosophical-Interdisciplinary Vanguard*, 5(3), 30-50.
<https://doi.org/10.26913/50302014.0112.0003>

- Miyazono, K., & Salice, A. (2020). Social epistemological conception of delusion. *Synthese*. <https://doi.org/10.1007/s11229-020-02863-1>
- Moore, A. T., & Schwitzgebel, E. (2018). The experience of reading. *Consciousness and Cognition*, 62, 57-68. <https://doi.org/10.1016/j.concog.2018.03.011>
- Moore, G. E. (1939/2013). Proof of an external world. In T. Baldwin (Ed.), *G. E. Moore: Selected writings* (pp. 147-170). Routledge. (Proceedings of the British Academy 25 (1939), pp.273-300)
- Moretti, L. (2020). *Seemings and epistemic justification: How appearances justify beliefs*. Springer.
- Mujica-Parodi, L. R., Malaspina, D., & Sackeim, H. A. (2000). Logical processing, affect, and delusional thought in schizophrenia. *Harvard Review of Psychiatry*, 8(2), 73-83. <https://doi.org/10.1080/hrp.8.2.73>
- Mullins, S., & Spence, S. A. (2003). Re-examining thought insertion. Semi-structured literature review and conceptual analysis. *British Journal of Psychiatry*, 182(04), 293-298. <https://doi.org/10.1192/bjp.182.4.293>
- Murphy, D. (2012). The folk epistemology of delusions. *Neuroethics*, 5(1), 19-22. <https://doi.org/10.1007/s12152-011-9125-5>
- Murphy, P., Bentall, R. P., Freeman, D., O'Rourke, S., & Hutton, P. (2018). The paranoia as defence model of persecutory delusions: A systematic review and meta-analysis. *The Lancet Psychiatry*, 5(11), 913-929. [https://doi.org/10.1016/s2215-0366\(18\)30339-0](https://doi.org/10.1016/s2215-0366(18)30339-0)
- Nadarevic, L., & Erdfelder, E. (2019). More evidence against the spinozan model: Cognitive load diminishes memory for "true" feedback. *Memory and Cognition*, 47(7), 1386-1400. <https://doi.org/10.3758/s13421-019-00940-6>

- Nasar, S. (1998). *A beautiful mind: The life of mathematical genius and nobel laureate john nash*. Simon & Schuster.
- Nelson, B., & Sass, L. A. (2017). Towards integrating phenomenology and neurocognition: Possible neurocognitive correlates of basic self-disturbance in schizophrenia. *Current Problems of Psychiatry*, 18(3), 184-200.
<https://doi.org/10.1515/cpp-2017-0015>
- Nelson, M. T. (2002). What justification could not be. *International Journal of Philosophical Studies*, 10(3), 265-281.
<https://doi.org/10.1080/09672550210152113>
- Nie, C. (2016). Delusional beliefs, two-factor theories, and bizarreness. *Frontiers of Philosophy in China*, 11(2), 263-278.
<https://doi.org/10.3868/s030-005-016-0020-1>
- Nie, C. (2017, 26 December 2017). When is a cognitive system immune to delusions? *Imperfect Cognitions*.
<http://imperfectcognitions.blogspot.com/2017/12/when-is-cognitive-system-immune-to.html>
- Nie, C. (2019). Continuing commentary: Challenges or misunderstandings? A defence of the two-factor theory against the challenges to its logic. *Cognitive Neuropsychiatry*, 24(4), 300-307.
<https://doi.org/10.1080/13546805.2019.1652156>
- Nie, C. (2021). Can a bodily theorist of pain speak mandarin? *Philosophia*. <https://doi.org/10.1007/s11406-021-00353-3>
- Noordhof, P., & Sullivan-Bissett, E. (2021). The clinical significance of anomalous experience in the explanation of monothematic delusions. *Synthese*. <https://doi.org/10.1007/s11229-021-03245-x>
- Nuara, A., Nicolini, Y., D'Orio, P., Cardinale, F., Rizzolatti, G., Avanzini, P., Fabbri-Destro, M., & De Marco, D. (2020). Catching the imposter in the brain: The case of capgras

- delusion. *Cortex*, 131, 295-304.
<https://doi.org/10.1016/j.cortex.2020.04.025>
- Owens, D. (2002). Epistemic akrasia. *Monist*, 85(3), 381-397.
<https://doi.org/10.5840/monist200285316>
- Oyebode, F. (2021). *Psychopathology of rare and unusual syndromes*. Cambridge University Press.
- Pacherie, E. (2008). Perception, emotions and delusions: The case of the capgras delusion. In T. Bayne & J. Fernández (Eds.), *Delusion and self-deception: Affective and motivational influences on belief formation* (pp. 107-125). Psychology Press.
- Pandis, C., Agrawal, N., & Poole, N. (2019). Capgras' delusion: A systematic review of 255 published cases. *Psychopathology*, 52(3), 161-173. <https://doi.org/10.1159/000500474>
- Parnas, J. (2004). Belief and pathology of self-awareness a phenomenological contribution to the classification of delusions. *Journal of Consciousness Studies*, 11(10-11), 148-161.
- Parnas, J., Bovet, P., & Zahavi, D. (2002). Schizophrenic autism: Clinical phenomenology and pathogenetic implications. *World Psychiatry*, 1(3), 131-136.
- Parnas, J., & Sass, L. A. (2001). Self, solipsism, and schizophrenic delusions. *Philosophy, Psychiatry, & Psychology*, 8(2), 101-120. <https://doi.org/10.1353/ppp.2001.0014>
- Parrott, M. (2016). Bayesian models, delusional beliefs, and epistemic possibilities. *The British Journal for the Philosophy of Science*, 67(1), 271-296.
<https://doi.org/10.1093/bjps/axu036>
- Parrott, M. (2021). Delusional predictions and explanations. *The British Journal for the Philosophy of Science*, 72(1), 325-353.
<https://doi.org/10.1093/bjps/axz003>

- Pascal, B. (1670/1995). *Pensées and other writings* (H. Levi, Trans.). Oxford University Press.
- Patterson, S. (2008). Clear and distinct perception. In J. Broughton & J. Carriero (Eds.), *A companion to descartes* (pp. 216-234). Blackwell. <https://doi.org/10.1002/9780470696439.ch13>
- Paul, E. S. (2020). Cartesian clarity. *Philosophers' Imprint*, 20(19), 1-28.
- Peters, M. A. K., Kentridge, R. W., Phillips, I., & Block, N. (2017). Does unconscious perception really exist? Continuing the assc20 debate. *Neuroscience of Consciousness*, 2017(1), Article nix015. <https://doi.org/10.1093/nc/nix015>
- Phillips, I., & Block, N. (2016). Debate on unconscious perception. In B. Nanay (Ed.), *Current controversies in philosophy of perception* (pp. 165-192). Routledge.
- Pienkos, E., Škodlar, B., & Sass, L. (2021). Expressing experience: The promise and perils of the phenomenological interview. *Phenomenology and the Cognitive Sciences*. <https://doi.org/10.1007/s11097-021-09731-4>
- Porcher, J. E. (2018). The doxastic status of delusion and the limits of folk psychology. In I. Hipólito, J. Gonçalves, & J. G. Pereira (Eds.), *Schizophrenia and common sense: Explaining the relation between madness and social values* (pp. 175-190). Springer. https://doi.org/10.1007/978-3-319-73993-9_11
- Porcher, J. E. (2019). Double bookkeeping and doxasticism about delusion. *Philosophy, Psychiatry, & Psychology*, 26(2), 11-119. <https://doi.org/10.1353/ppp.2019.0013>
- Prinz, J. J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford University Press.
- Pryor, J. (2000). The skeptic and the dogmatist. *Nous*, 34(4), 517-549. <https://doi.org/10.1111/0029-4624.00277>

- Pryor, J. (2004). What's wrong with moore's argument?
Philosophical Issues, 14(1), 349-378.
<https://doi.org/10.1111/j.1533-6077.2004.00034.x>
- Pryor, J. (2005). There is immediate justification. In M. Steup & E. Sosa (Eds.), *Contemporary debates in epistemology* (pp. 181-202). Blackwell.
- Qureshi, N., Al-Dossari, D., Al-Bedah, A., Salem, S., & Al-Habeeb, A. (2016). Delusional jealousy: A brief literature review and a case report of 'unfaithful husband' with sociocultural repercussions. *International Neuropsychiatric Disease Journal*, 6(2), 1-18. <https://doi.org/10.9734/indj/2016/22583>
- Radden, J. (2011). *On delusion*. Routledge.
- Raese, J., & Ibrahim, F. (2015). Psychosis and capgras delusions in a patient diagnosed with cogan's syndrome. *Clinical Case Reports and Reviews*, 1(1), 6-9.
<https://doi.org/10.15761/ccrr.1000103>
- Ramirez-Bermudez, J., Aguilar-Venegas, L. C., Crail-Melendez, D., Espinola-Nadurille, M., Nente, F., & Mendez, M. F. (2010). Cotard syndrome in neurological and psychiatric patients. *Journal of Neuropsychiatry and Clinical Neurosciences*, 22(4), 409-416. <https://doi.org/10.1176/appi.neuropsych.22.4.409>
- Ratcliffe, M. (2004). Interpreting delusions. *Phenomenology and the Cognitive Sciences*, 3(1), 25-48.
<https://doi.org/10.1023/b:Phen.0000041899.08813.38>
- Ratcliffe, M. (2008a). *Feelings of being: Phenomenology, psychiatry and the sense of reality*. Oxford University Press.
- Ratcliffe, M. (2008b). The phenomenological role of affect in the capgras delusion. *Continental Philosophy Review*, 41(2), 195-216. <https://doi.org/10.1007/s11007-008-9078-5>
- Ratcliffe, M. (2010). Delusional atmosphere and delusional belief. In S. Gallagher & D. Schmicking (Eds.), *Handbook of*

- phenomenology and cognitive science* (pp. 575–590).
Springer.
- Ratcliffe, M. (2017). *Real hallucinations: Psychiatric illness, intentionality, and the interpersonal world*. MIT Press.
- Ratcliffe, M., & Broome, M. R. (2018). *Beyond 'salience' and 'affordance': Understanding anomalous experiences of significant possibilities*.
- Raz, J. (2009). Reasons : Practical and adaptive. In D. Sobel & S. Wall (Eds.), *Reasons for action* (pp. 37–57). Cambridge University Press.
- Reed, E. J., Uddenberg, S., Suthaharan, P., Mathys, C. D., Taylor, J. R., Groman, S. M., & Corlett, P. R. (2020). Paranoia as a deficit in non-social belief updating. *eLife*, 9, Article e56345.
<https://doi.org/10.7554/eLife.56345>
- Reid, I., Young, A. W., & Hellowell, D. J. (1993). Voice recognition impairment in a blind capgras patient. *Behavioural Neurology*, 6(4), 225-228. <https://doi.org/10.3233/BEN-1993-6409>
- Reimer, M. (2009). Is the impostor hypothesis really so preposterous? Understanding the capgras experience. *Philosophical Psychology*, 22(6), 669-686.
<https://doi.org/10.1080/09515080903409945>
- Reimer, M. (2010). Only a philosopher or a madman: Impractical delusions in philosophy and psychiatry. *Philosophy, Psychiatry, & Psychology*, 17(4), 315-328.
<https://doi.org/10.1353/ppp.2010.0028>
- Renault, B., Signoret, J.-L., Debrulle, B., Breton, F., & Bolgert, F. (1989). Brain potentials reveal covert facial recognition in prosopagnosia. *Neuropsychologia*, 27(7), 905-912.
[https://doi.org/https://doi.org/10.1016/0028-3932\(89\)90066-3](https://doi.org/https://doi.org/10.1016/0028-3932(89)90066-3)
- Rey, G. (2020). The analytic/synthetic distinction. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2020 ed.).

<https://plato.stanford.edu/archives/win2020/entries/analytic-synthetic/>

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, 96(3), 538-558.

<https://doi.org/10.1037/a0014038>

Rinard, S. (2018). Believing for practical reasons. *Nous*, 53(4), 763-784. <https://doi.org/10.1111/nous.12253>

Rinard, S. (2019). Equal treatment for belief. *Philosophical Studies*, 176(7), 1923-1950. <https://doi.org/10.1007/s11098-018-1104-9>

Ritunnano, R., Humpston, C., & Broome, M. R. (2021). Finding order within the disorder: A case study exploring the meaningfulness of delusions. *BJPsych Bulletin*, 1-7.

<https://doi.org/10.1192/bjb.2020.151>

Rivolta, D., Palermo, R., & Schmalzl, L. (2013). What is overt and what is covert in congenital prosopagnosia? *Neuropsychology Review*, 23(2), 111-116.

<https://doi.org/10.1007/s11065-012-9223-0>

Roessler, J. (2013). Thought insertion, self-awareness, and rationality. In K. W. M. Fulford, M. Davies, R. Gipps, G. Graham, J. Sadler, T. Thornton, & G. Stanghellini (Eds.), *The oxford handbook of philosophy and psychiatry* (pp. 658-672). Oxford University Press.

Roessler, J. (2015). Thinking, inner speech, and self-awareness. *Review of Philosophy and Psychology*, 7(3), 541-557.

<https://doi.org/10.1007/s13164-015-0267-y>

Roessler, J. (2019). The manifest and the philosophical image of perceptual knowledge. In F. Stadler & C. Limbeck-Lilienau (Eds.), *The philosophy of perception: Proceedings of the 40th*

- international wittgenstein symposium* (pp. 275–302). de Gruyter. <https://doi.org/10.1515/9783110657920-016>
- Roessler, J. (forthcoming). Thought insertion and the ontology of thinking In P. López-Silva & T. McClelland (Eds.), *Intruders in the mind: Interdisciplinary perspectives on thought insertion*. Oxford University Press.
- Rojo, V. I., Caballero, L., Iruela, L. M., & Baca, E. (1991). Capgras' syndrome in a blind patient. *American Journal of Psychiatry*, 148(9), 1271-1272. <https://doi.org/10.1176/ajp.148.9.1271b>
- Rokeach, M. (1964/2011). *The three christs of ypsilanti: A psychological study*. New York Review Books.
- Rose, D., Buckwalter, W., & Turri, J. (2014). When words speak louder than actions: Delusion, belief, and the power of assertion. *Australasian Journal of Philosophy*, 92(4), 683-700. <https://doi.org/10.1080/00048402.2014.909859>
- Ross, R. M., McKay, R., Coltheart, M., & Langdon, R. (2015). Jumping to conclusions about the beads task? A meta-analysis of delusional ideation and data-gathering. *Schizophrenia Bulletin*, 41(5), 1183-1191. <https://doi.org/10.1093/schbul/sbu187>
- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, 126(11), 2381-2395. <https://doi.org/10.1093/brain/awg241>
- Ryle, G. (1949/2009). *The concept of mind*. Routledge.
- Sakakibara, E. (2018). Intensity of experience: Maher's theory of schizophrenic delusion revisited. *Neuroethics*, 12(2), 171-182. <https://doi.org/10.1007/s12152-018-9385-4>
- Sass, L. A. (1994). *The paradoxes of delusion: Wittgenstein, schreber, and the schizophrenic mind*. Cornell University Press.

- Sass, L. A. (2014). Delusions and double book-keeping. In T. Fuchs, T. Breyer, & C. Mundt (Eds.), *Karl jaspers' philosophy and psychopathology* (pp. 125-147). Springer.
- Sass, L. A., & Pienkos, E. (2013). Delusion: The phenomenological approach. In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, & T. Thornton (Eds.), *Oxford handbook of philosophy and psychiatry* (pp. 632-657). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199579563.013.0039>
- Scanlon, T. (1998). *What we owe to each other*. Harvard University Press.
- Scheibe, K. E., & Sarbin, T. R. (1965). Towards a theoretical conceptualisation of superstition. *The British journal for the philosophy of science*, 16(62), 143-158.
<https://doi.org/10.1093/bjps/XVI.62.143>
- Schweinberger, S. R., & Burton, A. M. (2003). Covert recognition and the neural system for face processing. *Cortex*, 39(1), 9-30.
[https://doi.org/https://doi.org/10.1016/S0010-9452\(08\)70071-6](https://doi.org/https://doi.org/10.1016/S0010-9452(08)70071-6)
- Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Nous*, 36(2), 249-275. <https://doi.org/10.1111/1468-0068.00370>
- Schwitzgebel, E. (2008). The unreliability of naive introspection. *The Philosophical Review*, 117(2), 245-273.
<https://doi.org/10.1215/00318108-2007-037>
- Schwitzgebel, E. (2012). Mad belief? *Neuroethics*, 5(1), 13-17.
<https://doi.org/10.1007/s12152-011-9127-3>
- Schwitzgebel, E. (2019). Belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 ed.).
<https://plato.stanford.edu/archives/fall2019/entries/belief/>
- Setiya, K. (2013). Epistemic agency: Some doubts. *Philosophical Issues*, 23(1), 179-198. <https://doi.org/10.1111/phis.12009>

- Sharp, H. M., Fear, C. F., Williams, J. M. G., Healy, D., Lowe, C. F., Yeadon, H., & Holden, R. (1996). Delusional phenomenology—dimensions of change. *Behaviour Research and Therapy*, 34(2), 123-142. [https://doi.org/10.1016/0005-7967\(95\)00059-3](https://doi.org/10.1016/0005-7967(95)00059-3)
- Siegel, S., & Silins, N. (2015). The epistemology of perception. In M. Matthen (Ed.), *Oxford handbook of philosophy of perception*. Oxford University Press.
- Signer, S. F., Van Ness, P. C., & Davis, R. J. (1990). Capgras's syndrome associated with sensory loss. *The Western journal of medicine*, 152(6), 719-720.
- Silva Jr, P., & Oliveira, L. R. G. (forthcoming). *Propositional and doxastic justification: New perspectives in epistemology*. Routledge.
- Simon, S. R., Khateb, A., Darque, A., Lazeyras, F., Mayer, E., & Pegna, A. J. (2011). When the brain remembers, but the patient doesn't: Converging fmri and eeg evidence for covert recognition in a case of prosopagnosia. *Cortex*, 47(7), 825-838. <https://doi.org/https://doi.org/10.1016/j.cortex.2010.07.009>
- Simonsen, A., Fusaroli, R., Petersen, M. L., Vermillet, A. Q., Bliksted, V., Mors, O., Roepstorff, A., & Campbell-Meiklejohn, D. (2021). Taking others into account: Combining directly experienced and indirect information in schizophrenia. *Brain*, 144(5), 1603-1614. <https://doi.org/10.1093/brain/awab065>
- Smithies, D. (2019). Epistemic akrasia. In *The epistemic role of consciousness* (pp. 284-311). Oxford University Press.
- So, S. H., Siu, N. Y., Wong, H. L., Chan, W., & Garety, P. A. (2016). 'Jumping to conclusions' data-gathering bias in psychosis and other psychiatric disorders - two meta-analyses of comparisons between patients and healthy individuals.

- Clinical Psychology Review*, 46, 151-167.
<https://doi.org/10.1016/j.cpr.2016.05.001>
- Southard, E. E. (1912). On the somatic sources of somatic delusions. *The Journal of Abnormal Psychology*, 7(5), 326-339. <https://doi.org/10.1037/h0070895>
- Spitzer, R. L., First, M. B., Kendler, K. S., & Stein, D. J. (1993). The reliability of three definitions of bizarre delusions. *The American Journal of Psychiatry*, 150(6), 880-884.
<https://doi.org/10.1176/ajp.150.6.880>
- Startup, M. (1997). Awareness of own and others' schizophrenic illness. *Schizophrenia Research*, 26(2), 203-211.
[https://doi.org/10.1016/S0920-9964\(97\)00050-9](https://doi.org/10.1016/S0920-9964(97)00050-9)
- Stephens, G. L., & Graham, G. (1994). Self-consciousness, mental agency, and the clinical psychopathology of thought insertion. *Philosophy, Psychiatry, & Psychology*, 1(1), 1-10.
- Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological Psychiatry*, 84(9), 634-643.
<https://doi.org/10.1016/j.biopsych.2018.05.015>
- Stompe, T., Ortwein-Swoboda, G., Ritter, K., & Schanda, H. (2003). Old wine in new bottles? Stability and plasticity of the contents of schizophrenic delusions. *Psychopathology*, 36(1), 6-12. <https://doi.org/10.1159/000069658>
- Stone, T., & Young, A. W. (1997). Delusions and brain injury: The philosophy and psychology of belief. *Mind & Language*, 12(3-4), 327-364. <https://doi.org/10.1111/j.1468-0017.1997.tb00077.x>
- Street, C. N. H., & Kingstone, A. (2017). Aligning spinoza with descartes: An informed cartesian account of the truth bias. *British Journal of Psychology*, 108(3), 453-466.
<https://doi.org/10.1111/bjop.12210>

- Sulik, J., Ross, R. M., Balzan, R., & McKay, R. (2021). Delusional ideation and data quality: Are classic cognitive biases an artefact of inattention? *PsyArXiv*.
<https://doi.org/10.31234/osf.io/ntsve>
- Sullivan-Bissett, E. (2020). Unimpaired abduction to alien abduction: Lessons on delusion formation. *Philosophical Psychology*, 33(5), 679-704.
<https://doi.org/10.1080/09515089.2020.1765324>
- Sunstein, C. R., & Vermeule, A. (2009). Conspiracy theories: Causes and cures. *Journal of Political Philosophy*, 17(2), 202-227.
<https://doi.org/10.1111/j.1467-9760.2008.00325.x>
- Szasz, T. S. (1974). *The myth of mental illness: Foundations of a theory of personal conduct* (Revised ed.). Perennial Library.
- Tanney, J. (2017). Trauma and belief. *Philosophy, Psychiatry, & Psychology*, 24(4), 351-353.
<https://doi.org/10.1353/ppp.2017.0051>
- Tenenbaum, S. (1999). The judgment of a weak will. *Philosophy and Phenomenological Research*, 59(4), 875-911.
<https://doi.org/10.2307/2653561>
- Teng, L. (2018). Is phenomenal force sufficient for immediate perceptual justification? *Synthese*, 195(2), 637-656.
<https://doi.org/10.1007/s11229-016-1233-7>
- Thornton, T. (2007). Why the idea of framework propositions cannot contribute to an understanding of delusions. *Phenomenology and the Cognitive Sciences*, 7(2), 159-175.
<https://doi.org/10.1007/s11097-007-9079-6>
- Todd, J. (1982). The history of the syndrome of capgras. In F. C. Rose & W. F. Bynum (Eds.), *Historical aspects of the neurosciences* (pp. 117-125). Raven Press.
- Todd, J., & Dewhurst, K. (1955). The othello syndrome: A study in the psychopathology of sexual jealousy. *Journal of Nervous and Mental Disease*, 122(4), 367-374.

- Tolhurst, W. (1998). Seemings. *American Philosophical Quarterly*, 35(3), 293-302.
- Tooming, U., & Miyazono, K. (2020). Vividness as a natural kind. *Synthese*. <https://doi.org/10.1007/s11229-020-02920-9>
- Torrìsi, M., De Luca, R., Pollicino, P., Leonardi, S., Marino, S., Maresca, G., Maggio, M. G., Piccolo, A., Bramanti, P., & Calabro, R. S. (2019). Poststroke delusions: What about the neuroanatomical and neurofunctional basis? *Applied Neuropsychology: Adult*, 26(4), 392-396. <https://doi.org/10.1080/23279095.2017.1421536>
- Tranel, D., & Damasio, A. (1985). Knowledge without awareness: An autonomic index of facial recognition by prosopagnosics. *Science*, 228(4706), 1453-1454. <https://doi.org/10.1126/science.4012303>
- Tranel, D., & Damasio, A. R. (1988). Non-conscious face recognition in patients with face agnosia. *Behavioural Brain Research*, 30(3), 235-249. [https://doi.org/10.1016/0166-4328\(88\)90166-0](https://doi.org/10.1016/0166-4328(88)90166-0)
- Tranel, D., Damasio, H., & Damasio, A. R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience*, 7(4), 425-432. <https://doi.org/10.1162/jocn.1995.7.4.425>
- Tucker, C. (2010). Why open-minded people should endorse dogmatism. *Philosophical Perspectives*, 24, 529-545. <https://doi.org/10.1111/j.1520-8583.2010.00202.x>
- Tucker, C. (2013a). Seemings and justification: An introduction. In C. Tucker (Ed.), *Seemings and justification: New essays on dogmatism and phenomenal conservatism* (pp. 1-29). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199899494.001.0001>

- Tucker, C. (Ed.). (2013b). *Seemings and justification: New essays on dogmatism and phenomenal conservatism*. Oxford University Press.
- Tumulty, M. (2011). Delusions and dispositionalism about belief. *Mind & Language*, 26(5), 596-628.
<https://doi.org/10.1111/j.1468-0017.2011.01432.x>
- Tumulty, M. (2012). Delusions and not-quite-beliefs. *Neuroethics*, 5(1), 29-37. <https://doi.org/10.1007/s12152-011-9126-4>
- Turner, M., & Coltheart, M. (2010). Confabulation and delusion: A common monitoring framework. *Cognitive Neuropsychiatry*, 15(1), 346-376. <https://doi.org/10.1080/13546800903441902>
- Turri, J. (2010). On the relationship between propositional and doxastic justification. *Philosophy and Phenomenological Research*, 80(2), 312-326. <https://doi.org/10.1111/j.1933-1592.2010.00331.x>
- Turri, J., Rose, D., & Buckwalter, W. (2018). Choosing and refusing: Doxastic voluntarism and folk psychology. *Philosophical Studies*, 175(10), 2507-2537. <https://doi.org/10.1007/s11098-017-0970-x>
- Uhlhaas, P. J., & Mishara, A. L. (2007). Perceptual anomalies in schizophrenia: Integrating phenomenology and cognitive neuroscience. *Schizophrenia Bulletin*, 33(1), 142-156.
<https://doi.org/10.1093/schbul/sbl047>
- Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, 79, 78-86.
<https://doi.org/10.1016/j.jesp.2018.06.010>
- Van Duppen, Z. (2015). The phenomenology of hypo- and hyperreality in psychopathology. *Phenomenology and the Cognitive Sciences*, 15(3), 423-441.
<https://doi.org/10.1007/s11097-015-9429-8>

- van Fraassen, B. C. (1989). *Laws and symmetry*. Oxford University Press.
- van Loon, M. (2021, September 9th-10th). *Delusions are responsive to reasons* Responsibility, Psychopathology & Stigma, Online. <https://www.neuroepigenethics.com/2021/07/30/conference-responsibility-psychopathology-stigma/>
- Vanderschraaf, P., & Sillari, G. (2021). Common knowledge. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 ed.). <https://plato.stanford.edu/archives/fall2021/entries/common-knowledge/>
- Voce, A., Calabria, B., Burns, R., Castle, D., & McKetin, R. (2019). A systematic review of the symptom profile and course of methamphetamine-associated psychosis substance use and misuse. *Substance Use & Misuse*, 54(4), 549-559. <https://doi.org/10.1080/10826084.2018.1521430>
- Von Domarus, E. (1944). The specific laws of logic in schizophrenia. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia* (pp. 104–114). University of California Press.
- Walker, C. (1991). Delusion: What did jaspers really say? *British Journal of Psychiatry*, 159(S14), 94-103. <https://doi.org/10.1192/s0007125000296566>
- Ward, T., & Garety, P. A. (2019). Fast and slow thinking in distressing delusions: A review of the literature and implications for targeted therapy. *Schizophrenia Research*, 203, 80-87. <https://doi.org/10.1016/j.schres.2017.08.045>
- Weatherson, B. (2008). Deontology and descartes's demon. *The Journal of Philosophy*, 105(9), 540-569. <https://doi.org/10.5840/jphil2008105932>
- Westermann, S., Gantenbein, V., Caspar, F., & Cavelti, M. (2018). Maintaining delusional beliefs to satisfy and protect

- psychological needs. *Zeitschrift für Psychologie*, 226(3), 197-203. <https://doi.org/10.1027/2151-2604/a000337>
- Wilkinson, S. (2013). Egocentric and encyclopedic doxastic states in delusions of misidentification. *Review of Philosophy and Psychology*, 4(2), 219-234. <https://doi.org/10.1007/s13164-012-0125-0>
- Wilkinson, S. (2016). A mental files approach to delusional misidentification. *Review of Philosophy and Psychology*, 7(2), 389-404. <https://doi.org/10.1007/s13164-015-0260-5>
- Williams, B. (1973). Deciding to believe. In B. Williams (Ed.), *Problems of the self: Philosophical papers 1956-1972* (pp. 136-151). Cambridge University Press.
- Williams, D. (2020). Socially adaptive belief. *Mind & Language*, 36(3), 333-354. <https://doi.org/10.1111/mila.12294>
- Williams, D., & Montagnese, M. (preprint). *Bayesian psychiatry and the social focus of delusions*. <http://philsci-archive.pitt.edu/18188/>
- Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press.
- Wittgenstein, L. (1969). *On certainty* (D. Paul & G. E. M. Anscombe, Trans.). Basil Blackwell.
- Woodward, J., & Hitchcock, C. (2003). Explanatory generalizations, part i: A counterfactual account. *Nous*, 37(1), 1-24. <https://doi.org/10.1111/1468-0068.00426>
- Young, A. W. (1998). *Face and mind*. Oxford University Press.
- Young, A. W. (2000). Wondrous strange: The neuropsychology of abnormal beliefs. *Mind & Language*, 15(1), 47-73. <https://doi.org/10.1111/1468-0017.00123>
- Young, A. W., & Ellis, H. D. (Eds.). (1989). *Handbook of research on face processing*. North Holland.
- Young, A. W., Hellawell, D., & De Haan, E. H. F. (1988). Cross-domain semantic priming in normal subjects and a

prosopagnosic patient. *The Quarterly Journal of Experimental Psychology Section A*, 40(3), 561-580.

<https://doi.org/10.1080/02724988843000087>

Young, A. W., & Leafhead, K. M. (1996). Betwixt life and death: Case studies of the cotard delusion. In P. W. Halligan & J. C. Marshall (Eds.), *Method in madness: Case studies in cognitive neuropsychiatry* (pp. 147-171). Lawrence Erlbaum Associates.

Young, A. W., Reid, I., Wright, S., & Hellowell, D. J. (1993). Face-processing impairments and the capgras delusion. *British Journal of Psychiatry*, 162(5), 695-698.

<https://doi.org/10.1192/bjp.162.5.695>

Young, G. (2007). Clarifying “familiarity”: Examining differences in the phenomenal experiences of patients suffering from prosopagnosia and capgras delusion. *Philosophy, Psychiatry, & Psychology*, 14(1), 29-37.

Young, G. (2008). Capgras delusion: An interactionist model. *Consciousness and Cognition*, 17(3), 863-876.

<https://doi.org/10.1016/j.concog.2008.01.006>

Young, G. (2009). In what sense 'familiar'? Examining experiential differences within pathologies of facial recognition. *Consciousness and Cognition*, 18(3), 628-638.

<https://doi.org/10.1016/j.concog.2009.06.006>

Yuksel, C., Yilmaz, S., Nesbit, A., Carkaxhiu, G., Ravichandran, C., Salvatore, P., Pingali, S., Cohen, B., & Ongur, D. (2018). Demographic and clinical characteristics associated with a history of bizarre delusions in a cross-diagnostic sample of individuals with psychotic disorders. *Asian Journal of Psychiatry*, 31, 82-85.

<https://doi.org/10.1016/j.ajp.2018.02.011>