

The clinical significance of anomalous experience in the explanation of monothematic delusions

Noordhof, Paul; Sullivan-Bissett, Ema

DOI:

[10.1007/s11229-021-03245-x](https://doi.org/10.1007/s11229-021-03245-x)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Noordhof, P & Sullivan-Bissett, E 2021, 'The clinical significance of anomalous experience in the explanation of monothematic delusions', *Synthese*, vol. 199, no. 3-4, pp. 10277-10309. <https://doi.org/10.1007/s11229-021-03245-x>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.



The clinical significance of anomalous experience in the explanation of monothematic delusions

Paul Noordhof¹ · Ema Sullivan-Bissett²

Received: 14 January 2020 / Accepted: 3 June 2021
© Crown 2021

Abstract

Monothematic delusions involve a single theme, and often occur in the absence of a more general delusional belief system. They are cognitively atypical insofar as they are said to be held in the absence of evidence, are resistant to correction, and have bizarre contents. *Empiricism* about delusions has it that anomalous experience is causally implicated in their formation, whilst rationalism has it that delusions result from top down malfunctions from which anomalous experiences can follow. Within empiricism, two approaches to the nature of the abnormality/abnormalities involved have been touted by philosophers and psychologists. One-factor approaches have it that monothematic delusions are a *normal* response to anomalous experiences whilst two-factor approaches seek to identify a clinically abnormal pattern of reasoning in addition to anomalous experience to explain the resultant delusion. In this paper we defend a one-factor approach. We begin by making clear what we mean by *atypical*, *abnormal*, and *factor*. We then identify the phenomenon of interest (monothematic delusion) and overview one and two-factor empiricism about its formation. We critically evaluate the cases for various second factors, and find them all wanting. In light of this we turn to our one-factor account, identifying two ways in which ‘normal response’ may be understood, and how this bears on the discussion of one-factor theories up until this point. We then conjecture that what is at stake is a certain view about the epistemic responsibility of subjects with delusions, and the role of experience, in the context of familiar psychodynamic features. After responding to two objections, we conclude that the onus is on two-factor theorists to show that the one-factor account is inadequate. Until then, the one-factor account ought to be understood as the default position for explaining monothematic delusion formation and retention. We don’t rule out the possibility that, for particular subjects with delusions there may be a second factor at work causally implicated in their delusory beliefs but, until the case for the inadequacy of the single factor is made, the second

This article belongs to the topical collection "Epistemological Issues in Neurodivergence and Atypical Cognition", edited by Alejandro Vázquez-del-Mercado Hernández and Claudia Lorena García-Aguilar.

Extended author information available on the last page of the article

factor is redundant and fails to pick out the minimum necessary for a monothematic delusion to be present.

Keywords Monothematic delusion · One-factor · Two-factor · Rationality · Epistemic responsibility · Normal range · Anomalous experience

1 Preliminaries

Delusional cognition is atypical, insofar as it uncommon among human believers. Indeed, its atypicality is part of its interest: why do only *some* folk hold such beliefs? We will see later that the two-factor theorist has motivated her theory by appeal to the atypicality of delusional belief within the set of folk with particular anomalous experiences (Sects. 3, 5). Atypicality is understood by appeal to what is abnormal. Where one-factor theorists only appeal to one abnormality (anomalous experience) in explaining delusional belief, two-factor theories claim we need to appeal to two (anomalous experience plus some bias, deficit, or performance error in mechanisms of belief production or evaluation). At the outset, we can distinguish two understandings of *abnormality*: statistical and functional.¹ The difference between them can be seen with the following example. Suppose that a particular reasoning style, *R*, occurs in all and only folk with delusions. Suppose also that *R* is *functionally normal*, which is to be understood as within the range of reasoning styles between which evolutionary selection has not distinguished. Two-factor theorists proceeding on the basis of statistical abnormality would take *R* to be a second abnormality, whilst those interested in functional abnormality would not. We take it that identifying a functional abnormality which afflicts those with delusions is more difficult than identifying a statistical one (since the former would require an independent account of functional normality). The literature has not been clear on the notion of normality at stake, but one reasonable interpretation is to understand two-factor theorists as seeking to identify a functional *abnormality* against a statistical assumption (viz. that these functional abnormalities are also statistically abnormal). Often some departure from rational standards of belief formation and evaluation is taken as a sign of functional abnormality. However, it is possible for subjects to be statistically abnormal in forming or evaluating beliefs by adhering more closely to these standards. Since two-factor theorists often object to one-factor theorists on the grounds that there is some irrationality in the formation or evaluation of the distinctive thematic beliefs of a subject with delusions, irrationality is a potentially distinct source of abnormality: involving significant departure from standards of rationality. We proceed by understanding the two-factor claim under the interpretation indicated with this qualification concerning the possibility of the second source.

¹ We mean the teleological/historical sense of ‘functional’ here, propounded by e.g. Ruth Millikan (1984), Karen Neander (1991), and others, and not e.g. Robert Cummins’s (1975) causal role sense or Christopher Boorse’s (1976) goal contribution sense.

Finally, we use the idea of a *factor* in the following way: a factor is something which is *different* between subjects with delusions and subjects without, something which is relevant to the project of explaining *delusional* belief in particular. It is for this reason that the second factor (or the way in which it is present) must be unique to folk with delusions; it is an abnormality that explains the formation of abnormal beliefs. This is something recognized by two-factor theorists. Martin Davies and colleagues for example take the second factor to be ‘a departure from what is normally the case’ (2005, p. 228); Ryan McKay and colleagues understand the deficit two-factor approach as one which ‘conceptualises delusions as involving dysfunction or disruption in ordinary cognitive processes’ (2010, pp. 316–317); and Tony Stone and Andrew Young talk of delusional reasoning being ‘abnormal’, and ‘differences between people with and without delusions’ (1997, p. 342). Where continuities with subjects without delusions are recognized, two-factor theorists still urge that subjects with monothematic delusions lie at the extreme end (McKay et al., 2010, pp. 316–317). We flag this option with our talk of the way in which the second factor is present.

The objective in the debate between one- and two-factor theorists is not to map every causal contribution to delusion formation. One-factor theories have been underestimated when this is forgotten. We do not deny that there might be particular reasoning styles involved in delusion formation; we only deny that such influences must be cognitive abnormalities. Where particular cases might involve abnormalities of this kind, it would be a mistake to generalize that finding to monothematic delusions simpliciter (Sect.4.4).

2 Monothematic delusion

Both of the most recent editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM IV, 2000, DSM V, 2013) define delusion in a way that makes it seem like a cognitive disorder: a belief is held ‘despite clear or reasonable contradictory evidence regarding its veracity’ (DSM V, 2013, p. 87). Philosophers and psychologists have similarly honed in on the supposed poor epistemic status of delusions, describing them as cases in which the belief is ‘unresponsive to considerations of plausibility and evidence’ (e.g. Davies & Davies, 2009, p. 285, cf. Flores, 2021).

Localized brain damage has been suggested as a source of the distinctive monothematic character of these delusions (Coltheart et al., 2007). Within a more complex delusory framework, it has seemed possible to focus upon particular elements of a subject’s belief system and look into the explanation for why *those* are present. We take up this framework for the purposes of discussion, though do not rule out the possibility that a successful treatment of monothematic delusions will extend to polythematic cases. We need not settle the matter here because discussion of monothematic delusions in particular originally motivated two-factor theories (e.g. Stone & Young, 1997, pp. 329–330).

The present debate between one- and two-factor theorists rests upon the assumption that monothematic delusions are beliefs. We are concerned with the question of how these beliefs are formed and maintained. Though the doxastic status of delusions is far from settled (for dissent see Currie, 2000; Egan, 2009; Dub, 2017), we

think that the case for doxasticism has been made persuasively elsewhere (Stone & Young, 1997, pp. 351–360; Bayne & Pacherie, 2005, Bortolotti, 2009, 2012).

Although delusions are conceived of as beliefs strongly resistant to counterevidence, most monothematic delusions are accompanied by some highly distinctive anomalous experiences which, themselves, constitute an, at least, apparent source of evidence in their own right for the beliefs in question. We will defend an empiricist account according to which the only clinically abnormal feature of subjects with delusions is their anomalous experiences. We deny that there need be a clinically abnormal feature at the cognitive level that explains the presence of monothematic delusions.

The table below gives various examples of monothematic delusions identifying the distinctive beliefs and hypothesized anomalous experiences at work.

Name	Belief	Experience
<i>Delusions of Misidentification</i>		
Capgras	An identified loved one has been replaced by an imposter	Lack of affective response to familiar people
Fregoli	An identified familiar person is disguised as others following me around	Affective response to unfamiliar people
Mirrored self-misidentification	The person I see reflected in the mirror is not me but somebody who looks very like me	Lack of experiencing the image in the mirror as oneself <i>or</i> mirror agnosia (taking the image in the mirror as presenting figures through a window)
<i>Delusions of self-consciousness</i>		
Cotard	I am dead	Lack of affective response, possibly generalised to environment
Somatoparaphrenia	Something which is, in fact, a part of my body (e.g. an arm) is not a part of my body	Lack of experience of limb as part of body image
Anosognosia	I can move a limb (when in fact I cannot)	Absence of experience of motoric failure
<i>Other delusions</i>		
Erotomania	An identified person is in love with me	?
Reverse othello	My lover has remained faithful to me (after e.g. abandonment)	?

Those delusions concerning which there is a question mark over the anomalous experience we return to at the end (Sect. 6).

3 One- and two-factor empiricism

All empiricists about monothematic delusion take the presence of anomalous experiences to be an important part of their explanation. Specific anomalous experiences provide an immediate explanation of their monothematic character (i.e. delusional beliefs are constrained to topics to which the anomalous experience is relevant).

Heightened affective responsiveness to most faces is plausibly an important part of an explanation of the development of Fregoli delusion, the belief that one is being followed by known but disguised people (Davies & Coltheart, 2000, pp. 31–32, Davies et al., 2001, p. 136). Absent affective responses in the experience of somebody one knows well is an important part of the explanation of the Capgras delusion that that person has been replaced by an imposter. One-factor accounts such as, most prominently, Brendan Maher's, hold that these experiences are the *sole clinically abnormal* cause of monothematic delusions, and that monothematic delusions are a normal response to such anomalous experiences (Maher, 1974, 1988, 1999, 2003, 2006).

Two-factor approaches are in disagreement with Maher's approach while building upon his insight. Whilst granting the importance of anomalous experience, they take a second factor to be required to explain the resultant delusion. All two-factor theorists argue that if a one-factor theory were true, then every subject who had the relevant anomalous experience would have the delusional belief. But this is not the case. So they reason that there must be some other causal factor involved (see e.g. Garety, 1991, p. 15; Garety et al., 1991, pp. 194–195, Chapman and Chapman, 1988, p. 174; Davies & Coltheart, 2000, pp. 11–12; Davies et al., 2001, p. 144; Young and De Pauw, 2002, p. 56; Davies et al., 2005, pp. 224–225; Coltheart, 2015, p. 23).

The claim that there are subjects without delusions who suffer the same anomalous experiences as those with delusions is contestable. Maher himself claimed that the kinds of experience subjects with delusions have are 'much more intense and prolonged' (Maher, 1999, p. 566), and are 'repeated or continue over an extended period' (Maher, 2006, p. 182). The thought is that if a subject really does have an anomalous experience of this sort, she will go on to develop the corresponding delusion.

Support for this is found in studies of ordinary subjects who undergo artificially induced anomalous experience in a laboratory setting, and show signs of delusion-like thinking in their experience reports. Maher refers to a study by Torsten Inge-mann Nielsen (1963b) in his discussion of this. Participants were given manipulated feedback about the movements they made with their hands, in some trials the feedback reflected the movements of the participants' hand, and in other trials it did not. In trials in which the hand seen behaved in ways which did not reflect the participants' movements, many participants explained the perceived discrepancy 'in terms strongly reminiscent of delusion explanations' (Maher, 2006, p. 182). Only two of twenty-eight participants guessed that an artificial hand was involved. Examples of delusion-like explanation include: 'It seems that my hand was moved by magnetism or electricity; I began to wonder if it was happening because I am homosexual; I was hypnotized; My hand was controlled by an outside physical force—I don't know what it was but I could feel it' (Nielsen, 1963b, cited in Maher, 2006, p. 182).² Maher takes from this that even brief and un-repeated anomalous experiences often

² The published work reporting on the study which Maher cites here does not include all of the experience reports (in particular, the one citing homosexuality). Nielsen published another paper in the same year in Danish, reporting on an earlier but similar experiment, with an English summary (1963a). We think that Maher may have had access to the Danish transcripts from one of these studies, which was not included in the English summary of the first study, nor in the English paper reporting on the second study.

prompt delusion-like explanations, which supports the claim that when such experiences are ‘more intense and prolonged’ (Maher, 1999, p. 566), they are sufficient for delusion formation. Putative cases of non-delusional subjects with the same anomalous experiences as those which gives rise to delusion in other subjects are in fact cases in which the experience is lacking in some respect (intensity, duration, etc.). It is these differences in the experiences which explains the difference between the delusional and non-delusional subjects, and not some further cognitive deficit, bias, or performance error characterizing only the former.

There is also debate over whether particular candidates of similar experiences at work are really ones in which this is so. For instance, Max Coltheart observes that the absence of autonomic responses taken to be indicative of anomalous experiences in Capgras subjects are also present in subjects with ventromedial lesions yet the latter do not suffer delusions (Tranel et al., 1995; Coltheart, 2007, pp. 1048–1049). However, as Sam Wilkinson has pointed out, the lesions are in different areas. Capgras patients tend to have right lateral temporal lesions and dorsolateral prefrontal damage rather than the ventromedial prefrontal damage observed by Daniel Tranel, Hanna Damasio, and Antonio R. Damasio (Wilkinson, 2015, p. 18, see Corlett, 2019 for further critical discussion of the Tranel, Damasio, and Damasio study and its implications for two-factor theories).

For the sake of argument, we will allow that there may be cases in which the anomalous experiences are appropriately similar. Our focus will be on whether two-factor theories are well-motivated even given this concession. It is important to recognize that nobody is claiming that the *sole causes* of delusional beliefs are anomalous experiences. There will be many causal factors along with the experiences both as part of the causal circumstances in which those experiences operate and causally downstream from the experience. Maher’s position would have been subject to easy refutation if his claim was that anomalous experiences are the sole cause. Already, in Maher’s talk of normal responses to anomalous experiences, a whole background of causal factors is taken into account. Two-factor theorists are claiming that, *against this assumed background*, an additional distinguishing factor needs to be acknowledged.

The envisaged second factor is a clinically significant abnormality relating to how delusional beliefs are formed or why they are retained (Langdon & Coltheart, 2000, pp. 201–206; Davies et al., 2001, p. 144). For example, and as noted earlier, Davies, Amiola Davies, and Coltheart talk of departures from what is normally the case (2005, p. 228). McKay, Robyn Langdon, and Coltheart talk of the second factor placing subjects at the extreme end of the belief evaluation continuum (2010, pp. 316–317). Others talk of ‘gross deviancy’ or abnormality (Chapman and Chapman, 1988, p. 179; Kaney & Bentall, 1989, p. 191; Stone & Young, 1997, p. 342). Everyday irrationalities are not properly so described. Sometimes two-factor theorists just talk of irrationalities in a way that would make their position no different from, but just a minor development of Maher’s position (e.g. Bortolotti, 2009, p. 57; Young and De Pauw, 2002, p. 56; Stone & Young, 1997, pp. 339–344). Although the latter still emphasize the ‘abnormality’ of the reasoning, this doesn’t seem to be their considered view. Two-factor theorists represent the current orthodoxy in combining an appeal to experience to account for the monothematic character of the delusion, while assuaging the intuition that there is some additional cognitive problem with the subject.

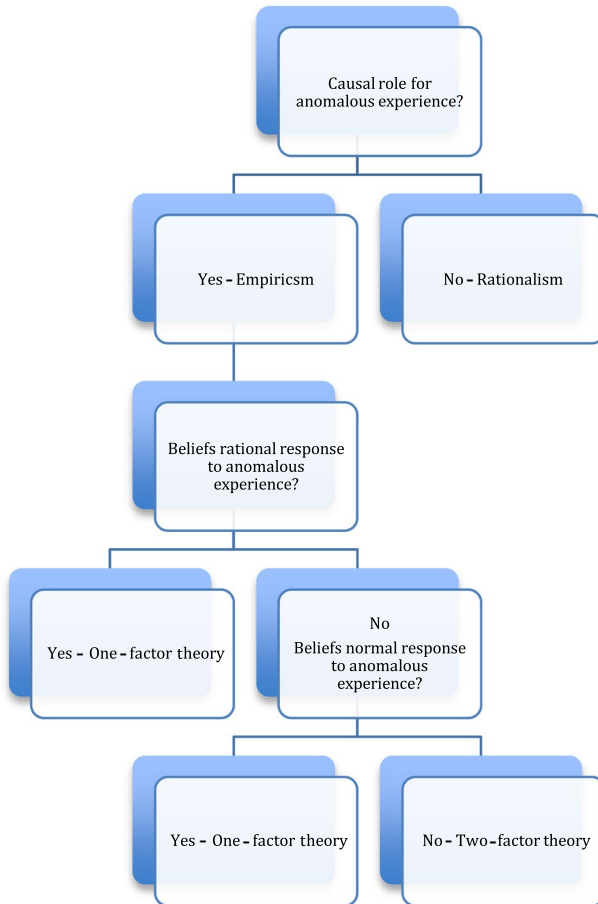
Talk of two causal factors in this framework involves a mix of type and token causal description that needs clarification. Two-factor theorists claim that there are two types of causal factor—individuated by location in the process of belief formation or evaluation—although they are neutral about whether they are the same or different kind of factor at these two locations. Corresponding to these two types of factor, the idea is that any subject with delusions will have tokens of the identified factors. These are anomalous experiences (first factor) and failures to form belief in the normal way, or evaluate it once formed, in the normal way (second factor). This distinction is important for the evaluation of the contribution of prediction-error theories of perception and cognition to this debate.

Prediction-error theories hold that perceptual processing involves the generation of predictions about sensory input, from antecedently held perceptual hypotheses about the world, with the aim of updating the hypotheses to minimize the error of these predictions on the basis of comparison between the predictions and sensory input. Delusions derive from the malfunctioning of this process, for example, faulty signals that a prediction isn't met leading to erroneous updating. The erroneous updating persists in the face of counterevidence because of continuing faulty signals supporting the updated hypothesis, mistaken downplaying of counterevidence due to the attribution of a high degree of noisiness to the incoming sensory input, or the relative absence of other kinds of sensory input bearing on the updated hypothesis (Corlett et al., 2007, p. 2396; Corlett et al., 2010, pp. 355–357; Corlett et al., 2016, p. 1146).

The candidate first factor—the contents of anomalous experiences—is usually said to be a result of either the combination of the antecedent hypothesis *plus* the prediction-not-met signal, the prediction-not-met signal giving, say, the sense of strangeness (or 'aberrant salience') in the experiences, *or* the reigning hypothesis which is the result of the signal (Kapur, 2003; Turner and Coltheart, 2010, p. 360; Hohwy, 2013, p. 48). To illustrate, in a case of the Capgras delusion, we either have the presentation of a partner with a sense that they are unfamiliar (the prediction of affective response is said to be not met) or *as* an imposter (see Bongiorno, 2020 for discussion on someone being presented *as an imposter* in experience).

Because the source of the error leading to delusional belief—malfunctioning of the prediction error process—is the same, whether it holds in the early stages of perceptual processing or in the later stages usually thought of as cognitive, we have the same type of factor in play. Prediction error accounts are one-factor accounts in this sense. The approach is neutral over the location(s) in which the error arises. Although its proponents often deny a sharp perception-cognition divide, they recognize the role of a more obviously cognitive place for a malfunctioning of this type. They conjecture that, in certain higher-level contexts, prediction error signals will be discounted (e.g. Corlett et al., 2016, p. 1148). Some two-factor theorists, thus, claim that prediction error theories are compatible with their approach and may help in its development (Coltheart, 2010, p. 25, see also Miyazono et al., 2014). So prediction error theories do not operate at the same level as the debate between one-factor and two-factor empiricism.

Here, in summary, is a sketch of the terrain so far.



We see that everything turns on the proper understanding of ‘normal response’. Sometimes ‘normal response’ is interpreted as the claim that the response to the experience is *rational*. The fact that there are some subjects who do not form the delusional belief upon the same anomalous experience is taken to indicate that those subjects who do are failing in their rationality. We consider later whether there is just one rational response to anomalous experience (Sect. 5.1), as well as developing the point that normal responses need not be rational ones (Sect. 5.2). There are irrationalities to which we are all prey. ‘Normal response’ in this extended sense means that subjects with delusions display no irrationality *beyond the normal range*. This second way of being a one-factor theorist has not been fully recognized in the debate.

4 Candidates for the second factor

Monothematic delusions ‘present in isolation in people whose beliefs are otherwise entirely unremarkable’ (Coltheart et al., 2007, p. 642). One-factor theorists find this segregation from other beliefs unsurprising, tracing it to the anomalous experiences to which the delusional beliefs are a response. Two-factor theorists face a difficulty. If subjects with these delusions have some clinically significant abnormality of their belief formation or evaluation processes, ‘why don’t they have a whole variety of delusional beliefs, rather than just one?’ (Coltheart, 2015, p. 25, see also Davies et al., 2001, pp. 149–150). Two-factor theorists must argue that while there is a clinically significant abnormality in belief formation or evaluation, this is not of such significance that it might give rise to delusional beliefs without persistent anomalous experiences. Some cognitive failing tips a subject into delusion (e.g. Langdon & Coltheart, 2000, p. 212).

The explanatory challenge for two-factor theories is that they need to provide a perspicuous characterization of the factor that leads to this tipping over. Here they face an unwelcome choice. If they emphasise that the second factor must be clinically significant, then the arguments they have offered in favour of a second factor are inadequate since all they cite in its support is that there is a variation in responses to anomalous experiences. We will explain how our preferred one-factor approach can accommodate such variation. On the other hand, if they merely highlight a particular second factor at work, which might be spread across the normal population, then their position is a development of Maher’s approach and not an alternative to it. Thus the self-image of two-factor theorists as departing from Maher will be shown to be either unmotivated or illusory.

With this task in mind, two-factor theories have come in, at least, three broad types: bias, deficit, and performance-failure theories (Bortolotti, 2009, p. 30). Sometimes theories are developed which run the first two together: there is a bias *because* there is a deficit (e.g. Coltheart et al., 2010, pp. 282–284). Biases are taken to be tendencies to depart from processing information in the way it should be processed. The relevant deficits are impairments to the normal system of belief formation, or subsequent evaluation, so that certain kinds of information cannot be processed (Bentall, 1995; Stone & Young, 1997, p. 331; Langdon & Coltheart, 2000, p. 202, Davies & Coltheart, 2000, p. 22). Performance-failure two-factor theorists deny that subjects lack the competence—thought of as a capacity—to form or evaluate beliefs unbiasedly (Gerrans, 2001). Instead, there is some kind of failure to put the competence into practice: a *performance* failure.

This categorization of the terrain raises two preliminary issues. The first is that the typology distinguishes theories that have different, not necessarily competing, aims. The most general characterization of the second factor is *some departure from a normal response regarding belief formation or evaluation*. Bias theories offer a more precise characterization of this departure. Deficit theories, on the other hand, are more concerned with what *causes* this departure. The corresponding causal story for bias theories is to talk of certain tendencies in the cognitive system. These tendencies may well be, but need not be, deficits. A bias may be present, indeed built

in, but is not the result of impairment. Performance failure accounts provide no further specification of the second factor but rather rule out a certain kind of characterization of it: one in terms of deficits or biases.

The second issue is whether a unitary account of the second factor is possible. For example, may some delusions arise from the operation of biases and others from deficits? The resolution of this issue has important consequences for our understanding of monothematic delusions. Consider two people with Capgras delusion. Subject A firmly holds onto the belief that his loved one has been replaced by an imposter in the face of considerable counterevidence but, eventually, abandons the belief. Subject B never does. She is absolutely resistant. Suppose that subject B has a deficit brought about by brain injury whereas subject A had no discernible brain damage. There are cases of this kind and, as we shall see, some cases of monothematic delusion—such as erotomania—rarely present brain damage (Sect. 4.2).

How should we view the two cases? One line might be that Subject A is not really a case of Capgras delusion but only an apparent one. Capgras delusion is a natural kind, with a hidden nature: the presence of a deficit. A second line would be that there are deficit and non-deficit instances of Capgras delusion. Capgras is variably realized. The second factor would be given a functional specification—a departure from a certain kind of normal relationship between experience and belief formation, or a failure of subsequent belief evaluation—and different realisations of this would be recognized. The presence of a deficit might be, but need not be, an explanation of the intractability of Subject B's condition. On the other hand, what was required for Capgras to present was something rather less than the postulated deficit.

When we look at the nature of, and evidence for, these three types of theories, we need to be clear about whether it supports the claim that a particular second factor is an *essential* part of monothematic delusion or, at best characterizes *some* cases and, at worst, fails to characterize what is required for a particular monothematic delusion.

4.1 Bias theories

Bias theories face two immediate difficulties. First, to characterize a bias away from how information should be processed, you have to provide some account of how it should be processed in the first place. Second, biases in the formation, and/or evaluation, of beliefs plausibly ought to apply across the board and not just to the distinctive topic of the relevant monothematic delusion. 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. By definition, though, it is a problem for monothematic delusions and the evidence for the presence of a bias here is correspondingly weaker.

To address the first issue, most bias theories are formulated within a Bayesian framework (e.g. Hemsley & Garety, 1986). To fix ideas, we shall adopt this framework and discuss the different candidate biases with Capgras delusion in mind.

Bayes' theorem holds:

$$P(h/e) = P(e/h) \cdot P(h)/P(e)$$

‘e’ is for evidence, ‘h’ for a hypothesis one is testing in the light of the evidence. The probabilities express degrees of confidence in a proposition rather than objective probabilities. My $P(\text{heads})$ may be 0.5 even if the coin is, in fact, biased and unbeknownst to me has an objective probability of turning up heads of 0.8. Suppose you are considering whether or not to believe h or not- h on the basis of evidence e . Then the following ratio captures which is more credible.

$$\frac{P(h/e)}{P(\text{not-}h/e)} = \frac{P(e/h) \cdot P(h)}{P(e/\text{not-}h) \cdot P(\text{not-}h)}$$

If $P(h/e)/P(\text{not-}h/e)$ is over 1, then $P(h/e)$ is more credible.

In the application of Bayes’ Theorem to delusional reasoning, we should take ‘h’ as a hypothesis about what is going on, and ‘e’ the anomalous experience. So, from the perspective of the subject with Capgras delusion, the probability attached to the hypothesis that *my partner has been replaced by an imposter*, given the subject’s suffering the relevant anomalous experiences, is the probability of the anomalous experience in which I have flat affect looking at my partner, given the delusional hypothesis, multiplied by the probability of my partner having been replaced by an imposter, divided by the probability of the anomalous experience.

To see whether it is rational for the subject to believe that their partner has been replaced by an imposter, we compare:

H_D : My partner has been replaced by an imposter

H_R : My partner remains as before

$P(H_D)$ will often be much smaller than $P(H_R)$. Ask a subject, before the onset of his delusion, ‘What’s the chance of your partner being replaced by an almost identical imposter?’ and he will say *very low*. Consider now $P(e/H_D)$ against $P(e/H_R)$. A subject experiencing no affective response to what seems to be his partner, given that she remains as before, is also very low. By contrast, it is considerably higher if a replacement has occurred.

Things may, then, be quite finely balanced to begin with (or favour H_R , McKay, 2012, pp. 339–341). This will depend upon the explanatory need if, say, the anomalous experiences seem very improbable. However, the imposter hypothesis increases in probability as the anomalous experiences build up. Within this framework, why do some subjects end up with Capgras delusion and others do not?

4.1.1 Jumping to conclusions/data gathering bias

Some argue that those with delusions suffer from a ‘jumping to conclusions’ bias (Garety, 1991; Garety et al., 1991). In the bead test, two jars, A and B, are filled with red and yellow beads in the following proportions: 85:15, 15:85 (Garety et al., 1991, p. 196); 60:40, 40:60 (Dudley et al., 1997a, pp. 252–256). Subjects are presented with beads, one at a time, without being told from which jar they are taken, and replaced, and they are asked to say when they are confident that the beads are coming from either jar A or jar B. For example, they might see red bead, red

bead, yellow bead, red bead, making jar A more likely. Subjects with delusions require fewer beads—hence the charge of ‘jumping to conclusions’— before saying whether they are from A or B. Likewise, the thought runs, Capgras subjects too quickly attach credibility to the hypothesis that their partner has been replaced by an imposter, given the anomalous experiences.

Further work has shown that a more precise specification of the ‘jumping to conclusions’ bias is that subjects with delusions have a ‘data gathering’ bias (Garety & Freeman, 1999, p. 131; Dudley et al., 1997a). Subjects with delusions display no difference from control subjects with respect to the estimation of probabilities of certain hypotheses on the basis of evidence as to frequency. They also display a similar way of balancing strength of evidence (e.g. all the beads being red in a small sample) over weight of evidence (a larger sample but some yellow beads). The difference is that subjects with delusions require fewer beads to be presented (less data) before they arrive at a conclusion (Dudley et al., 1997a, pp. 248–249); as do normal subjects rating high for delusional ideation (Linney et al., 1998, p. 299).

While there is evidence that this style of reasoning is present in schizophrenia detailed above, there is little evidence that it is present in monothematic delusion. Differences in reasoning between subjects with delusions and subjects without delusions are often found not to be statistically significant (e.g. McKay et al., 2007, pp. 368–369, Brakoulias et al., 2008, p. 157, pp. 161–162; Jacobsen et al., 2012, p. 12) [we have given an online page reference here because the only online version has not been updated to the journal pagination], or very slight (Kemp et al., 1997). The difficulties with the jumping to conclusions proposal don’t end there.

First, as its advocates note, the identified ‘jumping to conclusions’ reasoning style is not, in fact, a failure of rationality, characterized by the Bayesian formula. It turns out that, for subjects with delusions, as a group, the average number of beads required conforms closer to Bayesian rationality. Subjects without delusions on the other hand are ‘too cautious’ (Garety et al., 1991, p. 200; Garety, 1991; Maher, 2006, p. 180).

Some respond to this point by suggesting that the crucial observation is that subjects with delusions are *different*, from those without, even if, by the lights of Bayesian rationality, not irrational (e.g. Bentall 1994, pp. 346–347). While they might be different, such a position does not constitute grounds for postulating a second factor. Maher does not deny that there are differences, he only denies that there is a clinically significant second factor to identify. Two-factor theories standardly reject Maher’s position taking subjects with delusions to be either irrational or abnormal in the sense we have specified. If subjects with delusions prove to be more rational by Bayesian standards, then they are covered by Maher’s position. Moreover, while there might be some irrationalities—for example, overoptimistic self-appraisal as against depressive realism—for which evolution selects, being worse Bayesian reasoners is unlikely to be one of them. What benefits would accrue to such a subject, especially bearing in mind the difference is strongest regarding neutral subject matters (see the third point below)? In any event, it would be a mistake to characterize the fact that there is some difference as a *bias* since many are reasoning as they should unless, as Richard Bentall remarks, bias is taken to cover a tendency to ‘jump to conclusions’ *relative* to the general population regardless of whether that

tendency is rational. That's probably why other two-factor theorists deny that the second factor relates to belief formation as opposed to evaluation (Coltheart et al., 2010, pp. 277–278) (see Sect. 4.2).

Second, subjects who displayed a 'jumping to conclusions' bias showed no greater tendency to *stick* to a conclusion come what may. As one might expect, some work showed that they would jump to *other* conclusions on the basis of new information (Garety, 1991, p. 16, Garety et al., 1991, p. 200; Davies et al., 2001, p. 149). Other work found no greater tendency to stick to their conclusion than control subjects (Dudley et al., 1997a, p. 257). Those jumping to the conclusion that their partner has been replaced might be expected to jump to the opposite conclusion when faced with counterevidence such as the testimony of friends and other family members. This is not found. Moreover, evidence that subjects with delusions show no more resistance to changing their mind with regard to *neutral* topics shows that there can't be a general bias or fault in the *evaluation* of their beliefs.

Third, as remarked above, the principal *difference* between normal subjects and those prone to delusions is most often found in neutral topics (although not always reproduced, for example, Warman & Martin, 2006, p. 763; McKay et al., 2007, pp. 368–369, Brakoulias et al., 2008, p. 157, pp. 161–162, or very slightly so, Kemp et al., 1997). Since a significant number of subjects without delusions also have the 'jumping to conclusions' reasoning style, at best, we have a pre-disposing factor which is not clinically relevant in itself but may be significant when combined with anomalous experiences (Dudley et al., 2015, p. 656).

In the case of non-neutral topics, for example, concerning people like themselves, subjects with delusions and those without both proved to be more inclined to jump to conclusions, with no enhanced tendency to 'jump to conclusions' for those with delusions (Dudley et al., 1997b; Kemp et al., 1997). The delusional subject matter of the monothematic delusions we have been considering is non-neutral. So there is no reason to suppose that a 'jumping to conclusions' bias in subjects with delusions showing up in neutral topics is especially relevant in the explanation of their distinctive doxastic behaviour. There is *some* evidence that those prone to delusions are more likely to 'jump to conclusions' when the emotionally salient topic is negative (Warman & Martin, 2006, p. 763). In which case, the argument would have to be that this slight difference in emotionally salient negative cases, when combined with an anomalous experience, results in delusional beliefs. We would still be left with the puzzle of how this explains why some subjects fail to have delusions as opposed to a delay in the onset of delusions.

The combination of the failure to explain the *persistence* of delusions plus the relative smallness of the difference, *when* the difference primarily shows up, and how this could explain why some people fail to have delusions, throws into question the role of this reasoning bias.

4.1.2 Attributional styles

Subjects with delusions have also been said to have a distinct attributional style. Where depressed subjects look for explanations of events that find fault with themselves, non-depressed subjects with delusions, especially persecutory delusions,

externalize the source of the problem (Kaney & Bentall, 1989; Bentall 1994, pp. 347–353). Instead of identifying a problem with themselves in their relationship to their partner, they say that the partner is an imposter. This has been cited as an explanation for how a subject can switch between Capgras and Cotard delusions (Wright et al., 1993). The latter comes on when the subject is depressed.

Davies and colleagues argue that a problem with this proposal is that some subjects suffer both delusions *simultaneously* (2001, pp. 148–149; Joseph, 1986). However, not all the cases they cite support this charge. For example, RY particularly suffered from Cotard delusion when he awoke from sleep and was in a disorientated state, unable to distinguish vivid dreams from subsequent reality. This resolved through the day and the most persistent delusion he had was Capgras (Butler, 2000a). Geoffrey Wolfe and Kwame McKenzie's case report describes a man whose Capgras delusion resolved with anti-psychotic medication and developed the Cotard delusion *subsequently* due to hypothesized depression (Wolfe & McKenzie, 1994, p. 842).

Nevertheless the approach fails to have general application. While paranoid subjects display the externalizing attributional style, non-paranoid subjects with somatic delusions fail to have the corresponding internalizing attributional style in which such delusions might be rooted (Sharp et al., 1997, p. 5). There is also an alternative explanation of the cases that might be rooted in an externalizing attributional style, namely the observational bias discussed in the next section (Stone & Young, 1997, pp. 349–350).

Finally, attributional biases of all sorts—including the ones just identified—are spread across the normal population (Langdon & Coltheart, 2000, p. 198). These, and a subject's shifting moods, will have an influence on the beliefs formed. But they are not characterisations of a clinically significant abnormal second factor.

4.1.3 Observational bias over conservatism

In our initial discussion of the Capgras case, we compared the imposter hypothesis with the no replacement hypothesis. There is, of course, a third hypothesis:

H_C : I have Capgras delusion.

Why doesn't the Capgras subject consider this possibility? There is no evidence of a general hypothesis testing difficulty that might make a subject pass this option over (Bentall & Young, 1996, p. 374). An alternative is that the subject has a tendency to privilege apparent information revealed in experience—putative 'observational data'—to ground her beliefs rather than whether the beliefs count as the most minimal adjustment to the rest of her beliefs (the latter often dubbed *doxastic* or *epistemic conservatism* (e.g. Davies & Coltheart, 2000, pp. 16–17; McKay, 2012, pp. 343–344)).

Some proponents of this position have argued that the observational bias is, in fact, an explanatory bias, or that an additional explanatory bias should be recognized (McKay, 2012, pp. 343–344; Davies & Davies, 2009, p. 291). One reason offered derives from a certain way of developing empiricist theories of delusion. According to *endorsement* approaches, the content of the anomalous experience is pretty much identical to the content of the corresponding delusional belief, which, therefore, is

an endorsement of the content of the experience. *Explanationist* approaches take the content of the anomalous experience to fall short of the delusional belief. The belief involves an explanatory hypothesis as to why the anomalous experience takes the form it does (Bayne & Pacherie, 2004b, p. 82). The suggestion is that if explanatory approaches are the right way to develop the empiricist position, then the bias is correctly described as an explanatory one.

This is mistake for two reasons. First, the operation of a bias towards observational data does not require that the content of the delusional belief is presented in experience. An explanatory connection can be the basis of the central role given to observational data in the formation of belief. Second, H_C also explains the content of experience. So an emphasis on explanatory adequacy does not differentiate between those who adopt, and those who fail to endorse, the delusional hypothesis. The distinction that we need is a bias in favour of the information given in experience—the observational data—and not in favour of explanations of it.

More generally, it is questionable whether this bias, however it is described, covers the formation of delusional beliefs in the way envisaged. Consider a case of anosognosia where, for example, a subject denies that her left arm is paralyzed. There may be an absence of experienced motor failure to alert the subject to the issue. Nevertheless, there are also a whole range of experiences—seeing the arm inactive while still feeling sensation through it, failures to clap, and so on—where there is observational data crying out to be explained. If there is a bias at work here, it is not properly characterized as privileging what is observed (Davies et al., 2005, pp. 217–227). Equally, as Davies and Coltheart point out, if subjects with delusions had a straightforward bias in favour of observational adequacy so understood, we would expect them to be taken in far more frequently by visual and other illusions. There is no evidence that this is so (Davies & Coltheart, 2000, pp. 25–27).

4.1.4 Biases in general

Aside from concerns about the proper generality and specification of any bias, it is hard to understand how subjects with delusions' processing of information can be biased when they can acknowledge that others will find the beliefs at which they have arrived preposterous or far-fetched (e.g. Halligan et al., 1993, pp. 164–165). This suggests that their competence is untouched and, indeed, that the competence has no in-built biases (Gerrans, 2001, pp. 167–168). It might be thought some may only acknowledge that their beliefs are preposterous because they recall that others have told them they are. However, reports of their discussions indicate them grappling with the implausibility of the hypothesis which otherwise is suggested strongly by their experience. These subjects do not display biases elsewhere in their reasoning. In many ways, they reproduce the reflective evaluation of their beliefs that other subjects might engage in, and they are not surprised by the reception their belief receives. Taken together, this all suggests that they are not simply seeking to humour the questioner with recalled charges of implausibility but are rather getting there themselves. It is worth remembering that those conducting the interview discussions will not be challenging the subject with delusions (which is acknowledged to be counterproductive) but will rather be exploring their views about the delusory

hypothesis. Thus, if the subject with delusions were otherwise parroting remembered charges of implausibility, they might take the situation to present the opportunity to find a supporter in the interviewer. Indeed, this is such a common phenomenon that it has been identified as one of the two features of monothematic delusions with which any theory must come to grips (Davies et al., 2001, pp. 149–150).

Research into the potential biases behind monothematic delusion persists because it is observed that, so long as there is some such reasoning tendency at work, when subjects are deluded, it does not much matter if this falls within the normal range. Their identification is illuminating in itself. We don't contest that this may turn out to be so. But the research is better divorced from the two-factor programme. There is no reason to suppose that bias research will turn up a *single* second factor to characterize those subjects who form delusional beliefs as a result of anomalous experiences, nor that subjects with delusions will have more extreme instances of these biases than subjects in the normal range (as is often assumed, Ross et al., 2015, p. 1183; Chapman and Chapman, 1988, p. 179). Development of bias research within the two-factor programme tends to encourage the slide from more modest claims about biases to these stronger ones that is best avoided (e.g. Coltheart, 2007, p. 1059).

4.2 Deficit Theories

Proponents of deficit two-factor accounts have tended to focus not on belief formation but belief evaluation as the location of the problem (Langdon & Coltheart, 2000, pp. 204–213, Davies et al., 2001, pp. 149–153, although sometimes a memory deficit is postulated, Davies et al., 2005, pp. 230–231). Coltheart and colleagues hypothesise that the second factor is right hemisphere damage in the frontal lobe, which interferes with the belief evaluation mechanism (Coltheart, 2007, p. 1046, 1051; Coltheart et al., 2007, p. 644). Coltheart, Langdon, and McKay go so far as to say it will be present in all cases of monothematic delusion (Coltheart et al., 2007, p. 644). Subjects with delusions form their beliefs in line with Bayes' Theorem but they fail to update their beliefs, in light of evidence that they are false, as a result of the deficit. As Coltheart, Peter Menzies, and John Sutton put it, they fail to take the evidence as evidence upon which they should update their beliefs (Coltheart et al., 2010, p. 280).

A principal motivation offered for this position is that if there is only a reasoning bias, rather than deficit, one would expect that the reasons against a delusional belief would pile up until it is abandoned, whereas subjects with delusions are very resistant to abandoning their beliefs (Langdon & Coltheart, 2000, p. 202). Deficit theorists allow that our reasoning capacities may be unbiased and otherwise intact. Sometimes they just propose that the deficit is an abnormal feeling of conviction attached to certain beliefs (Turner and Coltheart, 2010, p. 363).

A recognised *prima facie* problem with deficit theories is that, as we have already discussed, subjects with monothematic delusions often acknowledge that the beliefs they have are implausible, which suggests they are able to evaluate them for plausibility, even while being unable to abandon them (Davies et al., 2001, p. 150;

Halligan et al., 1993, pp. 161–165). In which case, the information about the plausibility or otherwise of a belief must be able to be processed. It is just not acted upon.

A second problem touching on the motivation for deficit theories is that, as Coltheart, Langdon, and McKay later acknowledge, sufferers from Capgras delusion, and the like, *can* abandon their beliefs (Coltheart et al., 2007, p. 646; Turner and Coltheart, 2010, pp. 370–371). How does this happen if the belief evaluation system is damaged? The preliminary response is to point out that the fact that it is damaged does not mean that it fails to work *at all*. But then it is hard to see why we have a deficit rather than a bias in play in which evidence piles up and the delusion is eventually abandoned. Revision of beliefs often occurs in conjunction with medication. Perhaps those who propound deficits will rely upon this feature to distinguish their position from a bias.

A third problem is that one of the most resistant cases of monothematic delusion is not generally accompanied with evidence of damage. Primary erotomania potentially goes on for decades and yet no obvious neural damage has been found (Coltheart acknowledges that the neuropathological evidence is variable but not how it throws into question the motivation for postulating a deficit (2013, p. 107)). There are even documented cases of the Capgras delusion with no identifiable neural damage corresponding to the second factor, although the same absence of response to familiar faces (Brighetti et al., 2007). Obviously we don't deny that there will be differences in brain function but these observations throw into question whether the proper way to conceptualise the differences is in terms of a deficit resulting from damage. It underlines the point that the proper characterization of delusion is not at the level of deficit but rather what the deficit *realizes*, a certain departure from how information should be processed, whether inside or outside of the normal range.

If there is damage to the belief evaluation system and anomalous experience, then why is there variation in, for example, who is believed to be an imposter in the case of Capgras? (Berson, 1983, p. 971; Ellis & de Pauw, 1994, pp. 320–321; Brighetti et al., 2007, p. 191). Shouldn't all familiar faces be seen as imposters and a lot of other strange beliefs come along too? This suggests that there are additional motivational factors at work in a subject with delusions' response to anomalous experience (McKay et al., 2010, pp. 318–319; Turner and Coltheart, 2010, p. 366). Anomalous experience of a father towards whom one already has ambiguous feelings is more readily interpreted as of an imposter. Once it is recognized that there is an additional differentiating factor in play, the question arises whether it may play a far more significant role.

More recently, two-factor theorists have accepted that there may not be a deficit in every subject with a delusion (e.g. Coltheart, 2005, pp. 75–76, although they often insist that there will be an impairment in belief evaluation in any case of delusion (Coltheart, 2010, p. 23)). The reverse Othello case, erotomania discussed further below and also, plausibly alien abduction delusions, hypothesized by some to be based upon hypnagogic and hypnopompic hallucinations, are acknowledged to be cases where this may be so. Two-factor theorists suggest that in these cases, there might be a second factor at work—a tendency to find certain kinds of reasoning *attractive* involving strange causality, new age thinking etc.—which falls short of a

deficit (cf. Sullivan-Bissett, 2020). By our lights, this identifies a motivational feature that is better understood as an explanation of a performance failure.

4.3 Performance failure

The third kind of two-factor theory takes the second factor to be a performance failure. To illustrate the notion of performance failure in another area, Philip Gerrans cites failing to be able to speak, as a result of a stroke, as not a failure in one's linguistic competence but rather in the performance of it (Gerrans, 2001, p. 165).

If Bayesian belief updating is used to characterize what has been called *procedural rationality*, then the discussion so far has done little to show that there is a performance failure here. Gerrans looks to a wider notion of pragmatic rationality that covers the evaluation of the prior probabilities of hypotheses $P(e)$, $P(h)$, how one is probable given another, $P(e/h)$ and, one might reasonably assume, whether something should be taken as evidence at all. So, for example, those who suffer from Capgras display a performance failure of pragmatic rationality by taking the hypothesis that one's partner has been replaced by an imposter as more probable than it is (Gerrans, 2001, p. 164, pp. 167–168, other proponents of this position include Bayne & Pacherie, 2004a, p. 6).

The ground for taking subjects with delusions to have a performance failure, rather than a lack of competence, is that, since they recognize that their delusions are hard to believe, this is evidence that their competence is intact (e.g. JK, Young and Leafhead 1996, p. 158, discussed by Gerrans, 2001, pp. 167–171). The observation does not sit happily with the ascription of a performance failure either though. If subjects with delusions are capable of noticing that their belief is incredible, then this recognition is a successful performance of the competence (recall that a similar issue faced by bias theories, Sect. 4.1.4 and deficit theories, Sect. 4.2).

Moreover, characterizing the second factor in terms of performance failure is not an explanation of the second factor but just a description of the minimum that is presumed to occur by two-factor theorists in the case of delusion, namely that a subject is not forming beliefs in the way that they should. An explanation would account for *why* there was a performance failure.

We have already noted that there is an issue about why subjects with Capgras delusion don't take all those people with whom they have close relations and/or are familiar to be imposters. For example, although RY had Capgras delusion, he only had it with respect to his father with whom he had had trust issues prior to his accident and the onset of Capgras (Butler, 2000a, p. 686, other cases involve a feeling of ambivalence, Enoch & Trethowan, 1991, pp. 11–12). If this kind of 'psychodynamic' explanation is used to differentiate between those the subject does and does not believe are replaced by imposters, it is plausible that it is also relevant to explain the performance error in the first place. 'Performance error' is a placeholder for appeal to what may be quite standard individual differences—deriving from background motivations, beliefs, and context—that give rise to performance failings. On its face, it does not constitute an alternative account in itself.

Appeal to motivational factors can account for why delusional beliefs are held extremely firmly without appeal to a deficit. An example is B.X who came to suffer from reverse Othello syndrome after substantial head injury which, because of his disabilities, led to his abandonment by his romantic partner. He claimed that she had remained faithful to him and they had married. In this case, the delusion seems to be a protective strategy to deal with the calamitous setback B.X. had sustained. It was eventually abandoned as he recovered some degree of independence and received counter-evidence in the form of a starkly honest description of the situation (Butler, 2000b, pp. 86–88). The case nicely illustrates how a pattern of response can be irrational, and thus a performance error, without obviously being abnormal.

Other features may be responsible for the belief that a specified individual is an imposter that are entirely neutral on the question of whether a performance error has taken place. For example, prior to the onset of Capgras, there is often a period in which the subject hasn't seen the individual said to be replaced by an imposter and, sometimes, there are very slight but real changes in the individual. This, in addition, to a context that may be unfamiliar, or increased suspiciousness, can result in the delusion that an imposter has taken the place of a loved one (Fleminger, 1992, p. 298). Lower confidence levels may be set for beliefs in accord with this suspiciousness, but this still constitutes a response entirely within the normal range (Fleminger, 1992, p. 299).

Discussing with a subject their delusional beliefs can be a frustrating process. Consider JK who claimed that she was dead. When asked whether she could feel that her heart was beating, she said that given that she was dead, a beating heart was no sign of life, while acknowledging that other people would find this hard to believe (Young and Leafhead, 1996, p. 158). Nevertheless, it is not obvious that this resistance is any stronger than conversations one may have with somebody who is in an abusive relationship, acknowledges that the facts, as they stand, indicate the abusive character of his partner and yet fails to believe that this is the case because of some dimension of his first person experience (*'if you felt like me...'*) (see, for example, relevant discussion in Noordhof, 2003). This is a readily understood phenomenon—one which litters the course of human relationships—which is equally likely to be present when a normal subject is faced with anomalous experiences.

Some may be inclined to think that a cognitive second factor is inevitably in play in delusions. Consider cases in which anomalous experiences do not involve hallucinatory elements, but are characterised by a lack of something expected to be experienced, for example, lack of affect. In such cases it might be argued that they involve something cognitive—an experience failing to meet some expectation—which is a second factor. But notice that this putative second factor is not clinically anomalous in the reported cases. The expectation of an affective response when looking at our loved ones is appropriate. So it is important to keep separate that there may be something cognitive at work from there being something clinically abnormal.

4.4 Characterisation of particular cases

The criticisms against the various candidate second factors have taken a number of different forms. In some cases, for example the ‘jumping to conclusions’ bias, we have questioned whether they can play the envisaged explanatory role. In other cases, such as observational bias, we have questioned the claim to complete coverage of monothematic delusions. For deficits, there has been concern about motivation, coverage, and explanatory adequacy. All of these criticisms are compatible with particular subjects having some of these factors at work to explain particular features of their monothematic delusion. For example, the intractability of some subject’s monothematic delusion may be due to a severe deficit in their belief evaluation system. The fact that such a deficit could not explain how many subjects with monothematic delusions are capable of acknowledging the implausibility of their beliefs and revising their beliefs in certain circumstances does not vitiate the explanatory appeal to a deficit to explain the details of *this particular* case. Nothing we have argued rules that out. However, what is going on in a particular case should not be mistaken for providing insight into monothematic delusion as a distinctive kind of mental disorder. The candidate second factors we have identified don’t have the requisite generality and are not required to explain the distinctive features of such delusions. Focus on them distorts appreciation of the clinical significance of abnormal experience to which we now turn.

5 Outline of a one-factor theory

The experiential anomalies that subjects with delusions suffer are substantial (Maher, 2003, p. 18). In Capgras delusion, subjects experience a lack of affective response to faces, failing to experience someone with whom they are in a close relationship in the usual emotionally significant way. Evidence for this is the absence of any difference in autonomic response when experiencing those people familiar to them, and those who are not. This often seems to be a consequence of neural damage in the dorsal route between the visual cortex and the limbic system (Ellis & Young, 1990, p. 244). Similar damage seems present in those who suffer from Cotard delusion (Young et al., 1994). In Cotard delusion, though, it is hypothesized that subjects fail to experience any emotional feelings at all in response to their environment (Young et al., 1992, p. 800). One hypothesis, as yet not tested, is that the failure of autonomic response is generalised. The table we gave at the beginning gives further examples.

The anomalous experiences to which subjects with delusions are responding are typically taken to be conscious by one-factor and two-factor theorists alike, though they need not be (e.g. Davies et al., 2005, p. 222). More recently, Coltheart, Menzies, and Sutton have suggested that those with Capgras delusion are not conscious of the activities of their autonomic system and, hence, its failure to register a familiar person as such (Coltheart et al., 2010, p. 264; see also, for mirrored self misidentification Davies et al., 2005, p. 224).

One-factor theorists hold that a subject's delusional beliefs are normal attempts to form beliefs in the light of these experiences (Maher, 1988, p. 22). Thus Maher talks of the delusional belief being held 'because of evidence powerful enough to support it' (Maher, 1974, p. 99). He claims that delusional hypotheses are best thought of as like scientific theories—both 'serve the purpose of providing order and meaning for empirical data obtained by observation' (Maher, 1988, p. 20). Although subjects can hold delusional hypotheses in the face of counterevidence, this is analogous to what goes on in scientific theory change (Maher, 1974, p. 107). Even better theories can be resisted because they conflict with a scientist's commitment to her own theory. In the case of a subject with a delusion, to accept the theory proposed to supplant the subject's own theory would be 'tantamount to asking them to trust the evidence of other people's senses in preference to their own'; something which is 'not impossible' but 'not readily done by *most people*' (Maher, 1988, p. 25, our emphasis).

Subjects with delusions might have been in a state of considerable distress prior to adoption of the delusional hypothesis that explains why they are undergoing such experiences. The adoption may bring them some relief and attendant intellectual satisfaction at figuring things out (Mishara, 2010, p. 10). Normal subjects would be loath to give up such prizes. We need not attribute to subjects with delusions special resistance to having their beliefs challenged (Maher, 1974, p. 104; Maher, 2006, p. 182). In addition, everyday subjects engage in irrational—but non pathological—behaviour in ordinary life and scientific discovery when they do not give up their hypotheses easily (Maher, 1988, pp. 20–22).

Those with monothematic delusions will rightly take themselves to have experiences that are largely reliable but which keep supporting a non-standard delusional belief. These beliefs will, by their very nature, adjust a subject's assessment of the implausibility of the theory they support over time. The initial delusional beliefs are like a bridgehead for others.

Nothing in the statement of this approach suggests that anybody who has the definitive anomalous experiences must have the delusional belief as well. Two-factor theorists who take one-factor theorists to be committed in this way are making two assumptions. First, that for any particular anomalous experience, there is only *one* rational response to it. Second, that when Maher talks of delusional beliefs as a *normal response* to anomalous experiences, he means a *rational response* to the anomalous experiences (e.g. Bentall et al., 2001, p. 1149, Davies & Coltheart, 2000, p. 8, 12, Bortolotti, 2009, p. 57). For those two-factor theorists who place the second factor in the belief evaluation process, the subjects' response should not just be understood as initial response but rather initial response plus whatever occurs as a result of subsequent evaluation. Without these assumptions in place, the observation that subjects vary in the responses they make could not be taken to be an objection to a one-factor approach. A successful development and defence of a one-factor theory, then, has to explain how one or both of these assumptions are wrong.

5.1 Reasonable differences in response to experience

The first defence of a one-factor theory is to challenge the idea that there is one rational way in which to respond to anomalous experiences or, even if there is, take the disagreement on this matter to be reasonable. Suppose it seems to you that your partner has been replaced by an imposter (because of your lack of affective response). For any such, surprising, content about the way the world is, another hypothesis is that there is something wrong with you. Is it clear what rationality requires here?

Bias two-factor theorists are committed to a positive answer. But is this right? Consider the debate between epistemological sceptics and dogmatists about perceptual justification (Pryor, 2000). What is at stake here is the weight that should be given to experience in response to scepticism about the external world. Dogmatists claim that our *experiences* of objects and properties in the external world are an answer to sceptical concerns, insofar as they provide perceptual justification for our external world beliefs (e.g. Pryor, 2000, pp. 532–538). Perceptual experience bestows *prima facie* immediate (albeit defeasible) justification for perceptual beliefs, and *does not* give immediate *prima facie* justification for believing that one is in a sceptical scenario (Pryor, 2000, p. 536). By the same token, then, Capgras dogmatists experiencing the absence of affect generated by a loved one—where their experiences are otherwise reliable presentations of the world—may feel that they have reason to dismiss claims that, in fact, they have a psychological abnormality (that they are in the equivalent of a sceptical scenario where their experiences are misleading). Rather, they may feel that their experience of their loved one as unfamiliar justifies them in believing that they are in fact unfamiliar. In each case, the preferred hypothesis supported by experience faces defeaters. In the case of Capgras, this might take the form of the reception of that hypothesis from others. In the case of resistance to scepticism, the occurrence of dreams, hallucinations, illusions, and our putative inability to distinguish them from perception. All the phenomena that motivated scepticism in the first place.

Of course, just like in the debate between sceptics and dogmatists, some will prefer to give less weight to experience and more to theoretical considerations that may throw that experience into doubt. As one patient puts it, her mother looked different by ‘having the different *inside* of her [...] 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’ (Turner and Coltheart, 2010, p. 372).

In both cases, there is a debate about which is the rational response to the evidence of the senses. Perhaps both are rational responses or perhaps there is just a reasonable difference over which is the rational response although one or the other is correct. In both cases, a story can be told about why one should be sceptical about the content of one’s experiences (one might be a brain in a vat or dreaming in the first, one might be suffering from Capgras delusion in the second) but likewise both kinds of subjects may insist: I recognize I cannot rule out in a satisfactory way that I’m a brain envatted or an abnormal subject, but still these abnormal possibilities do not seem like real options to me (see Sullivan-Bissett, 2018 for discussion of how non-delusional alternative explanations might be unavailable to subjects with

delusions). As Maher says, such experiences ‘have the irreducible primary quality that sensory experiences generally have. They cannot be reasoned away’ (Maher, 2003, p. 18).

Note that the dogmatic response to scepticism doesn’t conflict with epistemic conservatism, the idea that the very holding of a belief provides some justification for that belief (McCain, 2008, p. 186). If one’s set of beliefs has justification just by being held, then one might be inclined to make only minimal adjustments to that set. While the challenge of scepticism threatens more than a minimal adjustment, it may seem otherwise with the case of monothematic delusion. It has been suggested, for example, by Brian McLaughlin that the belief that a loved one has been replaced by an imposter contradicts many of a person’s background beliefs (2009, p. 143).³ But this is overstated. Not only are works of non-fantastic fiction littered with examples of imposters taking the place of key figures, but also there is a rich vein of news reports or rumours about how people are replaced by look-alikes. As we noted earlier, particular cases of Capgras often involve a period during which the subject with the delusion hasn’t seen the individual believed to be an imposter so no magical event is supposed in which an imposter is substituted under the subject’s very eyes. The belief that one is faced with an imposter is a natural adjustment in response to the Capgras subject’s sensory experience. This is probably why two-factor theorists usually allow that the belief is a consequence of rational belief formation and place the irrationality in subsequent belief evaluation (e.g. Coltheart et al., 2010, pp. 277–278).

Even here, matters are not clear-cut. For one thing, subjects with such delusions are weighing the disturbance that the imposter hypothesis brings, along with increased suspicion about those who don’t accept it, against a view about how someone they believe they have feelings for, looks affectively to them. It is hard to underestimate how striking it is that someone looks emotionally different to you from the way you expect. As McLaughlin points out, when the delusory belief gets entrenched, epistemic conservatism may well assist its retention (McLaughlin, 2009, p. 150).

Equally, one of our central beliefs is that our own perceptual and cognitive processes are working effectively. If a subject has flat affect when experiencing her partner, the belief that *her partner has been replaced by an imposter* may seem like a radical change in her beliefs about the world. However, consider the alternative belief that *her perceptual and cognitive processes are not functioning properly*. Believing this would also represent a radical change. It is not obvious that epistemic conservatism favours the second belief over the first. After all, the alternative belief has potential implications for many of her other beliefs. It is by no means clear how isolated they can take this to be. The point is not that the radical change envisaged in the belief that one’s perceptual and cognitive processes are not functioning properly is scarier for the individual. It is pretty scary to imagine one’s loved one replaced by an imposter that everybody continues to recognize. The point concerns what is potentially more radical to one’s epistemic situation.

³ This point was also emphasised by an anonymous reviewer.

Of course a subject may entertain the hypothesis that they are suffering an entirely local malfunction that doesn't threaten radical revisions of their beliefs. Rejecting the imposter hypothesis does not require believing that there are widespread problems with their perceptual and cognitive processes. However, this more modest hypothesis has epistemological difficulties of its own and, from the subject's perspective, has substantial explanatory costs. If their experiences are otherwise reliable and don't reveal imposters everywhere, what grounds are there for supposing that a local malfunction is at work that the presentation of their senses might not be thought to outweigh? Likewise, the explanation of the anomalous experience provided by the imposter hypothesis is straightforward. They are faced with an imposter. By contrast, the hypothesis that one has Capgras delusion is a promissory note for an explanation which is, as yet, unclear (*why that loved one?*) put forward by an individual that they may not trust. In this respect, a more general malfunction of their perceptual and cognitive processes is more satisfying but is more radical in its implications.

Maybe these issues should be resolved one way rather than the other. But even if there is *one* weight we *should* give to experience, reasonable differences over what it is would both explain the variation in subjects' responses and yet make those responses reasonable (even if one of them is incorrect). So, even if you are convinced that rationality cannot support belief in the imposter hypothesis, this is not sufficient to undermine the first line of defence of a one-factor approach. There are reasonable differences over the rational response to Capgras experiences that makes Capgras subjects' responses reasonable if not rational.

5.2 Normal range irrationality

The second line of defence for a one-factor theory allows that the subject's response to their anomalous experiences may not be rational, but still holds that it is within the normal range of responses, concerning which we would not suppose that there is anything clinically amiss. Recall that a key motivation of two-factor theories is the observation that not all folk who undergo particular anomalous experiences become delusional. Thus there must be a second factor at work the presence of which explains why only some subjects adopt delusions in the face of such experiences. However, a one-factor theory can appeal to a normal range of responses to anomalous experiences which can explain the difference without positing a clinically significant cognitive contribution. The variation between those subjects who form, or persist in, delusional beliefs, and those who do not, is rather due to individual differences in intellectual style and character at work in *normal* belief formation.

On what grounds do we make this suggestion? On the grounds that even among the healthy population within which the first factor is not at work, there is a range of intellectual styles resulting in an array of strange beliefs. As Maher reminds us, *normal* members of the general population are 'prone to believe in the Bermuda Triangle, flying saucers, spoon-bending by mental power, the Abominable Snowman, and return to life after the out-of-body experience of death' as well as 'prebirth hypnotic age regression, multiple personalities, [...] and so forth' (Maher, 1988, p. 26).

We add that there are ‘at least several thousand worldwide’ (French et al., 2008, p. 1387) who believe they have been abducted by aliens. Key to this bizarre belief is that researchers interested in explaining it have *not* sought a cognitive abnormality, even whilst recognizing that there are people who have the abduction experience without the abduction belief. Rather, researchers have instead sought to identify a range of normal range cognitive styles that may help explain why some folk form the abduction belief and others do not (see Sullivan-Bissett, 2020 for discussion). There is no reason, we say, not to extend this methodology to monothematic delusions in general, but doing so is a very different thing from seeking to identify a cognitive contribution understood as a second factor.

Obviously the line between normal irrationality and the kind of abnormality that requires postulation of a second factor is going to be hard to draw. This is not a problem that we face and the two-factor theorists avoid. Indeed, the burden is greater upon them since the second factor—a bias, deficit, or performance error—is identified as responsible for an abnormal response that differentiates those with delusions.

5.3 Epistemic responsibility

It would be better for the whole debate if we could settle on a characterization of normal range reasoning and, thereby, a normal range response, and its abnormal contrast case. A natural way of thinking about the issue is to take the relevant dimension of abnormality to involve mental processes that are *significantly more resistant to rational correction* than we would expect, in general, from the population, *given the presence in that population of the kind of experiences which subjects with delusions undergo*. There are two elements to this idea (italicized above).

Normal irrationalities are those over which the subject has some responsibility the basis of which shows up in typical patterns of responsiveness to reasons. Normal subjects are irresponsible if they are insufficiently cautious in their reasoning for a belief. They are in possession of reasons—as the contents of other beliefs—or have had reasons presented to them which, had they considered them properly, would have led to the abandonment of the belief. Those whose irrationality is *abnormal* are not just epistemically irresponsible in this way but are outside the realm of epistemic responsibility. Reflection on the considerations for their beliefs would not help them in abandoning them, or stymie their formation in the first place.

The first element links the characterisation of normal versus abnormal irrationalities to some of the things virtue epistemologists like Linda Zagzebski say. The epistemic vices of intellectual pride and conformity are things for which we are responsible and can exercise control (Zagzebski, 1996, p. 60, see also Cassam 2015a, 2015b).

Nevertheless, the second element is that it is undeniable that some kinds of experience may place our capacities to respond to reasons under significant pressure due to their disturbing character. So we need to adjust our expectations of rational responsiveness to the kind of experiences subjects with monothematic delusions are undergoing. That is the second element.

Emphasising that subjects with delusions are, at worst, only normally irrational is to claim that their epistemic irresponsibility in the face of abnormal experiences results in delusional beliefs. There is nothing further wrong with them. It is possible to appeal to their reasoning powers, with some hope of success, to a comparable degree to debates we might enter with those prone to conspiracy theories or committed to views what about they have undergone in the workplace that might turn them into vexatious litigants. The only difference is the disturbing character of the experiences people with delusions are undergoing.

There is evidence to support this understanding of subjects with monothematic delusions. For example, LU was challenged in her Cotard belief that she was dead by pointing out that she was still capable of motion whereas dead people she had observed were not. In this case, her delusion resolved and she abandoned her belief (McKay & Ciolotti, 2007, p. 353). Similarly, MF—who had Capgras delusion—became convinced that his wife was, in fact, his wife when he was challenged to find another explanation for the fact she was wearing their wedding ring with her initials inscribed upon it (Coltheart, 2007, p. 1054; Coltheart et al., 2007, p. 646, for another example, see Turner and Coltheart, 2010, p. 371). Studies showing that subjects with delusions respond to cognitive behavioural therapy, permanently recording a reduction in credence to their delusions are also relevant here (e.g. Brakoulias et al., 2008, pp. 161–163). It has also been observed that those sliding into a delusion of misidentification, or coming out of it, closely inquire into the nature and background of the suspected double showing sensitivity to rational factors in belief formation and later evaluation (Christodoulou, 1978, p. 70).

That does not mean that reasoning with any subject with a monothematic delusion can result in them abandoning the delusional belief. Far from it! Our point is only that this feature of delusional beliefs (their seeming irrevisability, or the epistemic irresponsibility manifest when failing to revise them) is both not unique to delusions and those holding them, and it is also understandable given the anomalous experiences subjects with delusions are undergoing. That subjects with delusions are reluctant to give up their beliefs even in the face of disconfirming evidence, that they are epistemically irresponsible, is characteristic of a whole range of normal range beliefs. Take self-deceivers (see Noordhof, 2003 for discussion) and vexatious litigants: discussing the beliefs of these folk can be just as unrewarding an experience and yet they fall within the normal range of subjects.

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). Belief in conspiracy theories is widespread, for example, a recent poll found that 63% of those registered to vote in the US buy into one or more conspiracy theories (Fairleigh Dickinson University

poll, 2013). But 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. Drawing on the empirical work on conspiratorial ideation, Joseph M. Pierre notes that it is ‘essentially a normal phenomenon’:

Indeed, many of the cognitive biases and other psychological quirks that have been found to be associated with [belief in conspiracy theories] are universal, continuously distributed traits varying in quantity as opposed to all-or-none variables or distinct symptoms of mental illness. They are present in those who do not believe in conspiracy theories and some of them, like need for uniqueness or closure, may be valued or adaptive in certain culturally-mediated settings. (Pierre, 2020, p. 618, see Douglas et al., 2019 for a review of the literature on the genesis of conspiratorial beliefs.)

When explaining the formation and maintenance of conspiratorial beliefs then, it has been recognized that it is a mistake to seek to identify a single cognitive contribution (abnormal or otherwise) as responsible, and talk of biases in this context is not used to pick out deviation from normal range, but rather departure from rationality. Those biases hypothesized to play a role are ones found in the general, non-conspiratorial, population. As Cassam puts it, ‘[t]here is no single or simple explanation of conspiracy mindedness, but there was never any serious hope of that. The answer to the question of why people believe in conspiracy theories is: it’s complicated’ (Cassam, 2019, pp. 61–62). Similar things can be said for folk with monothematic delusions. It is a mistake of the two-factor theorist to look for a single cognitive feature (let alone a clinically abnormal one) to do the work of explaining monothematic delusion formation. Rather, the epistemic irresponsibility displayed by such subjects is representative of normal range irrationality, which, when understood against a background of distressing anomalous experiences, is both understandable and not all that surprising.

6 Two objections

One objection often raised is that, if it is found that in all cases of monothematic delusion, or a certain type of monothematic delusion, there is brain damage in a certain part of the brain, then this is compelling evidence that a second factor is at work.

The observation of brain damage does not establish whether we have a clinically abnormal deficit of the relevant kind. Even if there are consequences, these may be no more than shifting a subject from one kind of normal response to another. By analogy, brain damage that results in a change in personality, doesn’t necessarily result in an abnormal personality but just a different point in the normal range of personalities.

We don't deny that, if, contrary to what we have argued, there did turn out to be some common cognitive deficit, this would be of explanatory interest. Just as we would not deny that if somebody with a certain body type were especially prone to a certain type of disease, this would be of interest. The issue is the clinical significance of this. Suppose the cognitive deficit were less pronounced in many subjects with monothematic delusions than shows up in conspiracy theorists and vexatious litigants. The former suffer highly anomalous experiences to which they are seeking to respond rather than facing more mundane features of life. It would not be appropriate to suppose that the candidate cognitive factor was of particular importance, even whilst granting its salience, in understanding monothematic delusions. The research orientation of two-factor theorists is otherwise. That is the point of taking the second factor to be clinically significant, putting such emphasis on it, and it is this that we say, as yet, there are no good grounds.

A second objection to the one-factor approach stems from monothematic delusions in which there is no obvious anomalous experience (those marked with a '?' in the experience column of our table presented earlier). Primary erotomania has an explosive onset and long term focus (the condition can last for decades) on a particular high status individual who is believed to be secretly in love with the sufferer, the belief being often highly resistant to alteration (Jordan & Howe, 1980, p. 980, pp. 984–985; Jordan et al., 2006, pp. 790–791). Ordinary experiences are interpreted in special ways, for example, the license plates of cars of a certain type from a particular state or the colour purple were taken as messages of love (Jordan et al., 2006, p. 788). In this case, the charge runs, there must be a cognitive deficit and if here, then why not in other cases of monothematic delusion?

A weakness in this objection is that the cognitive aspect to it is strikingly different, both in its explosive rather than gradual onset, and in its generation of the delusion without the support of anomalous experience to provide the content. Those more inclined to see continuities with other monothematic delusions conjecture that such cases may result out of an erroneous attribution of salience to events in experience (Coltheart, 2010, pp. 24–25). If the latter is correct, perhaps supported by motivational factors, then we have the basis for a one-factor approach to erotomania.

Such cases suggest two ways in which one-factor theories may be developed further. First, primary erotomania often arises after some emotional trauma such as divorce, with a previous history of difficult relationships with family members of an older generation of the opposite sex, a sense of relative social isolation, and suspiciousness. This may plausibly generate a motivational framework in which certain combinations of everyday experiences are misinterpreted, resulting in future experiences of events being given inappropriate salience that serve to enmesh a subject further in delusory beliefs. The distinctive tie to *anomalous* experience is loosened.

Second, the recurrent talk of suspiciousness suggests an extension of the notion of experience to mental modules which display many of its distinctive features, especially that of relative informational encapsulation: that is, other beliefs don't tend to influence or moderate what is presented to be the case. Thus, Joel Gold and Ian Gold have proposed that a malfunctioning suspicion system, which enables subjects to detect threatening intent, is at the root of many delusions (Gold & Gold, 2014, pp.

191–197). Erotomania is suggested to be a motivated response to absence of social status and the consequent threat perceived as a result of this.

7 Conclusion

Two-factor approaches have proven attractive because they capture the role of experience but recognize the intuitive thought that delusions involve some abnormal cognitive failing. Nevertheless, as we have seen, the arguments in their favour don't work. The second factor has proven extremely elusive and may prove neither abnormal nor unified. By contrast, one-factor theories provide a plausible and unified treatment of the nature of monothematic delusions. As a result, the one-factor approach should be treated as the default hypothesis. Unless there is substantial theoretical reason from the study of subjects with delusions to postulate a second factor it should be the way we approach understanding such subjects.

One-factor theories also offer an important difference of emphasis. They see subjects with delusions as, to an extent, victims of experience and motivational factors. Anomalous experiences can be particularly hard to resist. However, at the end, we saw how experience does not have to be anomalous to result in a delusion. Rather a sequence of experiences that unlocks a motivational generation of a particular belief, can then result in anomalous interpretations of experiences which are presented as obvious further support for the belief. The debate between one-factor and two-factor theories is not then, simply, a debate about classification and demarcation. It is a debate over delusional entrapment and the role of experience in shaping our mental lives. These are dangers we are otherwise prone to underestimate by loose talk of a second factor.

Acknowledgements With thanks to the Arts and Humanities Research Council for funding the research of which this paper is a part (*Deluded by Experience*, grant no. AH/T013486/1). We are grateful for comments on some of the material in this paper presented to audiences in Durham, Gargnano, and St. Andrews, as well as to two reviewers for this journal.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- American Psychiatric Association (2000), (2013) Diagnostic statistical manual of mental disorders, Fourth, Fifth edition.
- Bayne, T., & Pacherie, E. (2004a). Bottom-up or top-down? Campbell's rationalist account of monothematic delusions. *Philosophy, Psychiatry, and Psychology*, 11(1), 1–11.

- Bayne, T., & Pacherie, E. (2004b). 'Experience belief, and the interpretative fold. *Philosophy, Psychiatry, and Psychology*, 11(1), 81–86.
- Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusion. *Mind and Language*, 20(2), 163–188.
- Bentall, R. P. (1994) Cognitive biases and abnormal beliefs: Towards a model of persecutory delusions. In: Anthony David, John C. Cutting (eds.) *The Neuropsychology of Schizophrenia*. Lawrence Erlbaum Associates, Publishers. 337–360.
- Bentall, R. P. (1995). Brains, biases, deficits and disorders. *British Journal of Psychiatry*, 167, 153–155.
- Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., & Kinderman, P. (2001). Persecutory delusions: A review and theoretical integration. *Clinical Psychology Review*, 21(8), 1143–1192.
- Bentall, R. P., & Young, H. F. (1996). Sensible hypothesis testing in deluded, depressed and normal subjects. *British Journal of Psychiatry*, 168, 372–375.
- Berson, R. J. (1983). Capgras' syndrome. *American Journal of Psychiatry*, 140(8), 969–978.
- Bongiorno, F. (2020). Is the Capgras delusion an endorsement of experience? *Mind & Language*, 35(3), 293–312.
- Borse, C. (1976). Wright on functions. *The Philosophical Review*, 85(1), 70–86.
- Bortolotti, L. (2009). *Delusions and Other Irrational Beliefs*. Oxford University Press.
- Bortolotti, L. (2012). In defence of modest doxasticism about delusions. *Neuroethics*, 5, 39–53.
- Brakoulias, V., Langdon, R., Sloss, G., Coltheart, M., Meares, R., & Anthony, H. (2008). Delusions and reasoning: A study involving cognitive behavioural therapy. *Cognitive Neuropsychiatry*, 13(2), 148–165.
- Brighetti, G. B., Paola, B., & Rosita, and Ottaviani, Cristina. (2007). "Far from the heart far from the eye": Evidence from the Capgras delusion. *Cognitive Neuropsychiatry*, 12(3), 189–197.
- Butler, P. V. (2000a). Diurnal variation in Cotard's syndrome (copresent with Capgras delusion) following traumatic brain injury. *Australasian and New Zealand Journal of Psychiatry*, 34, 684–687.
- Butler, P. V. (2000b). Reverse Othello syndrome subsequent to traumatic brain injury. *Psychiatry*, 63(1), 85–92.
- Cassam, Quassim. (2015a). 'Bad thinkers'. Aeon. Published 13th Mar 2015. Access 15th Feb 2016. URL: <https://aeon.co/essays/the-intellectual-character-of-conspiracy-theorists>.
- Cassam, Q. (2015b). Stealthy Vices. *Social Epistemology Review and Reply Collective.*, 4(10), 19–25.
- Cassam, Q. (2019). *Conspiracy Theories*. Polity Press.
- Chapman L. J., Chapman J. P. (1988) 'The genesis of delusions'. In: T. F. Oltmanns, B. A. Maher (eds.) *Delusional Beliefs* (pp. 167–183). Wiley.
- Christodoulou, G. N. (1978). Course and prognosis of the syndrome of doubles. *The Journal of Nervous and Mental Disease*, 166(1), 68–72.
- Coltheart, M. (2005). Delusional belief. *Australian Journal of Psychology*, 57(2), 72–76.
- Coltheart, M. (2007). Cognitive neuropsychology and delusional belief. *The Quarterly Journal of Experimental Psychology*, 60(8), 1041–1062.
- Coltheart, M. (2010). The neuropsychology of delusions. *Annals of the New York Academy of Sciences*, 1191, 16–26.
- Coltheart, M. (2013). On the distinction between monothematic and polythematic delusions. *Mind and Language*, 28(1), 103–112.
- Coltheart, M. (2015). From the internal lexicon to delusional belief. *AVANT*, 3, 19–29.
- Coltheart, M., Langdon, R., & McKay, R. (2007). Schizophrenia and monothematic delusions. *Schizophrenia Bulletin*, 33(3), 642–647.
- Coltheart, M., Menzies, P., & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive Neuropsychiatry*, 15(1), 261–287.
- 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.
- 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.
- Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R. F., 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, 2387–2400.
- Corlett, P. R., Taylor, J. R., Wang, X.-J., Fletcher, P. C., & Krystal, J. H. (2010). Towards a Neurobiology of delusions. *Progress in Neurobiology*, 92(3), 345–369.
- Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, 72(20), 741–765.
- Currie, G. (2000). Imagination, delusion and hallucinations. *Mind and Language*, 15(1), 168–183.

- Davies, A. M., & Davies, M. (2009). Explaining pathologies of belief. In L. Bortolotti & M. Broome (Eds.), *Psychiatry as Cognitive Neuroscience* (pp. 285–323). Oxford University Press.
- Davies, M., Anne, D. A., & Coltheart, M. (2005). Anosognosia and the two-factor theory of delusions. *Mind and Language*, *20*(2), 209–236.
- Davies, M., & Coltheart, M. (2000). Introduction: Pathologies of belief. *Mind and Language*, *15*(1), 1–46.
- Davies, M., Coltheart, M., Langdon, R., & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry, and Psychology*, *8*(2–3), 133–158.
- de Ellis, H. D., & Pauw, Karel. (1994). The cognitive neuropsychiatric origins of Capgras delusion. In S. D. Anthony & C. C. John (Eds.), *The Neuropsychology of Schizophrenia* (pp. 317–335). Lawrence Erlbaum Associates.
- Douglas, K., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding conspiracy theories. *Political Psychology*, *40*(1), 3–35.
- Dub, R. (2017). Delusions, acceptances, and Cognitive feelings. *Philosophy and Phenomenological Research*, *94*(1), 27–60.
- Dudley, R. E. J., John, C. H., Young, A. W., & Over, D. E. (1997a). Normal and abnormal reasoning in People with Delusions. *British Journal of Clinical Psychology*, *36*, 243–258.
- Dudley, R. E. J., John, C. H., Young, A. W., & Over, D. E. (1997b). The effect of self-referent material on the reasoning of people with delusions. *British Journal of Clinical Psychology*, *36*, 575–584.
- Dudley, R. E. J., Taylor, P., Wickham, S., & Hutton, P. (2015). ‘Psychosis, delusions and the ‘jumping to conclusions’ reasoning bias: A systematic review and meta-analysis. *Schizophrenia Bulletin*, *42*(3), 652–665.
- Egan, A. (2009). Imagination, delusion, and self-deception. In T. Bayne & J. Fernández (Eds.), *Delusion and Self-Deception* (pp. 263–280). Psychology Press.
- Ellis, H. D., & Young, A. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry*, *157*, 239–248.
- Enoch, M. D., & Trethowan, W. (1991). *Uncommon Psychiatric Syndromes* (Third Edition). Oxford, Heinemann Ltd.
- Fairleigh Dickinson University’s Public Mind Poll 2013: <http://publicmind.fdu.edu/2013/outthere/final.pf>
- Fleminger, S. (1992). Seeing is believing: The role of ‘Preconscious’ perceptual processing in delusional misidentification. *British Journal of Psychiatry*, *160*, 293–303.
- Flores, C. (2021). Delusional evidence-responsiveness. *Synthese*. <https://doi.org/10.1007/s11229-021-03070-2>
- French, C. C., Sanromauro, J., Hamilton, V., Fox, R., & Thalbourne, M. A. (2008). Psychological aspects of the alien contact experience. *Cortex*, *4*, 1387–1395.
- Garety, P. (1991). Reasoning and delusions. *British Journal of Psychiatry*, *159*(14), 14–19.
- Garety, P. A., & Freeman, D. (1999). Cognitive Approaches to delusions: A critical review of theories and evidence. *British Journal of Clinical Psychology*, *38*, 113–154.
- Garety, P. A., Hemsley, D. R., & Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients: Biases in performance on a probabilistic inference task. *The Journal of Nervous and Mental Disease*, *179*(4), 194–201.
- Gerrans, P. (2001). Delusions as performance failures. *Cognitive Neuropsychiatry*, *6*(3), 161–173.
- Gold, J., & Gold, I. (2014). *Suspicious Minds*. Free 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*, 159–166.
- Hemsley, D. R., & Garety, P. A. (1986). The formation and maintenance of delusions; a Bayesian analysis. *British Journal of Psychiatry*, *149*, 51–56.
- Hohwy, J. (2013). *The Predictive Mind*. Oxford University Press.
- Jacobsen, Pamela, Freeman, Daniel, & Salkovskis, Paul. (2012). Reasoning bias and belief conviction in obsessive-compulsive disorder and delusions: Jumping to conclusions across disorders? *British Journal of Clinical Psychology*, *51*(1), 84–99.
- Jordan, H. W., & Howe, G. (1980). De Clerambault syndrome (Erotomania): A review and case presentation. *Journal of the National Medical Association*, *72*(10), 979–985.
- Jordan, H. W., Lockert, E. W., Johnson-Warren, M., Cabell, C., Cooke, T., Greer, W., & Howe, G. (2006). Erotomania revisited: Thirty four years later. *Journal of the National Medical Association*, *98*(5), 787–793.

- Joseph, A. B. (1986). Cotard's syndrome in a patient with coexistent capgras' syndrome, syndrome of subjective doubles and palinopsia. *Journal of Clinical Psychiatry*, *47*(12), 605–606.
- Kaney, S., & Bentall, R. P. (1989). Persecutory delusions and attributional style. *British Journal of Medical Psychology*, *62*, 191–198.
- 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.
- Keeley, B. (1999). Of conspiracy theories. *The Journal of Philosophy*, *96*(3), 109–126.
- Kemp, R., Chua, S., McKenna, P., & Anthony, D. (1997). Reasoning and delusions. *British Journal of Psychiatry*, *170*, 398–405.
- Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind and Language*, *15*(1), 184–218.
- Linney, Y. M., Peters, E. R., & Ayton, P. (1998). Reasoning biases in delusion-prone individuals. *British Journal of Clinical Psychology*, *37*, 285–302.
- Maher, B. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology*, *30*(1), 98–113.
- Maher, B. (1988). Anomalous experience and delusional thinking: The logic of explanations. In Thomas Oltmanns & Brendan Maher (Eds.), *Delusional Beliefs*. John Wiley and Sons.
- Maher, B. (1999). Anomalous experience in everyday life: Its significance for psychopathology. *The Monist*, *82*(4), 547–570.
- Maher, B. (2003). Schizophrenia, aberrant utterance and delusions of control: The disconnection of speech and thought, and the connection of experience of and belief. *Mind and Language*, *18*(1), 1–22.
- Maher, B. (2006). The relationship between delusions and hallucinations. *Current Psychiatry Reports*, *8*, 179–183.
- Mandik, P. (2007). Shit happens. *Episteme*, *4*(2), 205–218.
- McCain, K. (2008). The virtues of epistemic conservatism. *Synthese*, *164*(2), 185–200.
- McKay, R. (2012). Delusional inference. *Mind and Language*, *27*(3), 330–355.
- McKay, R., & Ciolotti, L. (2007). Attributional style in a case of Cotard delusion. *Consciousness and Cognition*, *16*, 349–359.
- McKay, R., Langdon, R., & Coltheart, M. (2007). Jumping to delusions? Paranoia, probabilistic reasoning and the need for closure. *Cognitive Neuropsychiatry*, *12*(4), 362–376.
- McKay, R., Langdon, R., & Coltheart, M. (2010). “Sleights of mind”: Delusions, defences and self-deception. *Cognitive Neuropsychiatry*, *10*(4), 305–326.
- McLaughlin, B. (2009). Monothematic delusions and existential feelings. In T. Bayne & J. Fernández (Eds.), *Delusion and Self-Deception* (pp. 139–164). Psychology Press.
- Millikan, R. (1984). *Language, Thought and Other Biological Categories*. MIT Press.
- Mishara, A. L. (2010). Klaus Conrad (1905–1961): Klaus Conrad (1905–1961): Delusional mood, psychosis, and beginning schizophrenia. *Schizophrenia Bulletin*, *36*(1), 9–13.
- Miyazono, K., Bortolotti, L., & Broome, M. (2014). Prediction-error and two-factor theories of delusion formation: Competitors or allies. In N. Gailbraith (Ed.), *Aberrant Beliefs and Reasoning* (pp. 34–54). Psychology Press.
- Neander, K. (1991). The teleological notion of “function.” *Australasian Journal of Philosophy*, *69*(4), 454–468.
- Nielsen, T. I. (1963a). Forstyrrelser og paradokser I viljelivet—Fremkaldt eksperimentelt. *Nordisk Psykologi*, *15*(3), 183–192.
- Nielsen, T. I. (1963b). Volition: A new experimental approach. *Scandinavian Journal of Psychology*, *4*, 225–230.
- Noordhof, P. (2003). Self-deception, interpretation and consciousness. *Philosophy and Phenomenological Research*, *67*(1), 75–100.
- Pierre, J. M. (2020). Mistrust and misinformation: A two-component, socio-epistemic model of belief in conspiracy theories. *Journal of Social and Political Psychology*, *8*(2), 617–641.
- Pryor, J. (2000). The skeptic and the dogmatist. *Noûs*, *34*(4), 517–549.
- 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.
- Sharp, H. M., Fear, C. F., & Healy, D. (1997). Attributional style and delusions: An investigation based on delusional content. *European Psychiatry*, *12*, 1–7.

- Stone, T., & Young, A. W. (1997). Delusions and brain injury: The philosophy and psychology of belief. *Mind and Language*, 12(3–4), 327–364.
- Sullivan-Bissett, E. (2018). Monothematic delusion: A case of innocence from experience. *Philosophical Psychology*, 31(6), 920–947.
- Sullivan-Bissett, E. (2020). Unimpaired abduction to alien abduction: Lessons on delusion formation. *Philosophical Psychology*, 33(5), 679–704.
- Tranel, D., Damasio, H., & Damasio, A. R. (1995). Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience*, 7(4), 425–432.
- Turner, Martha, & Coltheart, Max. (2010). Confabulation and delusion: A common monitoring framework. *Cognitive Neuropsychiatry*, 15(1–3), 346–376.
- Warman, D. M., & Martin, J. M. (2006). Jumping to conclusions and delusion proneness: The impact of emotionally salient stimuli. *The Journal of Nervous and Mental Disease*, 194(10), 760–765.
- Wilkinson, S. (2015). Delusions, dreams and the nature of identification. *Philosophical Psychology*, 28(2), 203–226.
- Wolfe, G., & McKenzie, K. (1994). Capgras, Fregoli and Cotard's syndromes and Koro in folie à deux. *British Journal of Psychiatry*, 165(6), 842.
- Wright, S., Young, A. W., & Hellowell, D. J. (1993). Sequential Cotard and Capgras delusions. *British Journal of Clinical Psychology*, 32, 345–349.
- Young, A. W., & Leafhead, K. (1996). Betwixt life and death: Case studies of the Cotard delusion. In P. Halligan & J. Marshall (Eds.), *Methods in Madness*. Psychology Press.
- Young, A. W., Leafhead, K. M., & Szulecka, T. K. (1994). The Capgras and Cotard delusions. *Psychopathology*, 27, 226–231.
- Young, A., Pauw, De., & Karel, W. (2002). One stage is not enough. *Philosophy, Psychiatry and Psychology*, 9(1), 55–59.
- Young, A. W., Robertson, I. H., Hellowell, D. J., de Pauw, K. W., & Pentland, B. (1992). Cotard delusion after brain injury. *Psychological Medicine*, 22, 799–804.
- Zagzebski, L. T. (1996). *Virtues of the Mind*. Cambridge University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Paul Noordhof¹ · Ema Sullivan-Bissett²

✉ Ema Sullivan-Bissett
e.l.sullivan-bissett@bham.ac.uk

Paul Noordhof
paul.noordhof@york.ac.uk

¹ Department of Philosophy, University of York, York YO10 5DD, UK

² Department of Philosophy, University of Birmingham, Birmingham B15 2TT, UK