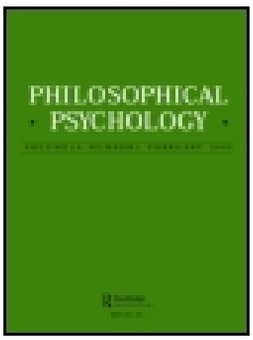
Monothematic Delusion:

Sullivan-Bissett, Ema

DOI:
10.1080/09515089.2018.1468024

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Sullivan-Bissett, E 2018, 'Monothematic Delusion: A Case of Innocence from Experience', *Philosophical Psychology*. https://doi.org/10.1080/09515089.2018.1468024



Philosophical Psychology

ISSN: 0951-5089 (Print) 1465-394X (Online) Journal homepage: http://www.tandfonline.com/loi/cphp20

Monothematic delusion: A case of innocence from experience

Ema Sullivan-Bissett

To cite this article: Ema Sullivan-Bissett (2018): Monothematic delusion: A case of innocence from experience, Philosophical Psychology, DOI: 10.1080/09515089.2018.1468024

To link to this article: https://doi.org/10.1080/09515089.2018.1468024

Published online: 23 May 2018.

PHILOSOPHICAL PSYCHOLOGY, 2018
https://doi.org/10.1080/09515089.2018.1468024

Routledge
Taylor & Francis Group

Monothematic delusion: A case of innocence from experience

Ema Sullivan-Bissett

Department of Philosophy, University of Birmingham, Birmingham, UK

ABSTRACT

Empiricists about monothematic delusion formation agree that anomalous experience is a factor in the formation of these attitudes, but disagree markedly on which further factors (if any) need to be specified. I argue that epistemic innocence may be a unifying feature of monothematic delusions, insofar as a judgment of epistemic innocence to this class of attitudes is one that opposing empiricist accounts can make. The notion of epistemic innocence allows us to tell a richer story when investigating the epistemic status of monothematic delusions, one which resists the trade-off view of pragmatic benefits and epistemic costs. Though monothematic delusions are often characterized by appeal to their epistemic costs, they can play a positive epistemic role, and this is a surprising conclusion on which, so I argue, all empiricists can agree. Thus, I show that all empiricists have the notion of epistemic innocence at their disposal.

ARTICLE HISTORY
Received 29 November 2014
Accepted 10 January 2018

KEYWORDS
Anomalous experience; delusion; delusion formation; epistemic innocence; monothematic delusion; one-factor; prediction error; two-factor

1. Introduction

The notion of epistemic innocence is gaining currency in the philosophical and psychological literature. It is used to capture the status of an epistemically faulty cognition which nonetheless has epistemic benefits. Lisa Bortolotti has discussed motivated delusions (2015) and elaborated and systematized delusions in schizophrenia (2016). Chris Letheby focuses on psychedelic states (2015), Katherine Puddifoot on inaccurate social cognition (2017), and in other work I have considered explanations of decisions or actions guided by implicit biases (Sullivan-Bissett, 2015) and clinical memory distortions (Bortolotti and Sullivan-Bissett, forthcoming). These authors argue that tokens of their target phenomenon are epistemically innocent.

Here, I am not extending the concept of epistemic innocence to another set of psychological phenomena. It is interesting and important to subject various

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

imperfect cognitions to such an analysis, since this will have implications for practices of epistemic evaluation. But once one does that, the question I am interested in is, to whom would it matter? Put another way: assume that the ascription of epistemic innocence to all of the states listed above is well motivated and plausible, which theorists is this verdict open to? Might some accounts of the nature of these phenomena block the arguments for their epistemic innocence from getting started? For example, if a non-doxastic account of delusions is correct, do the arguments given by Bortolotti on the epistemic innocence of some delusions (2015, 2016) still apply? If one's account of confabulation excludes motivational factors, can one accept my (2015) claim that some confabulatory explanations are epistemically innocent? Or, to bring us to the question of this paper: if some monothematic delusions are epistemically innocent, is this something which all empiricists about their formation—one- and two-factor theorists alike—can accommodate? If some monothematic delusions are epistemically innocent, this would matter to our understanding of them in general, and we would need to know how this would integrate into current debates concerning delusion formation. Although I will speak to whether some monothematic delusions are epistemically innocent, my primary interest here is in determining whether such a judgment is available to different theorists of monothematic delusion formation. If it is, we have found in this notion something which can unify competing accounts of monothematic delusion formation.

Why should we care about epistemic innocence in this context? Appealing to it would give us a novel account of the epistemic status of delusions, allowing for a more balanced view of their role in the lives of the subjects who have them. The notion of epistemic innocence demonstrates the importance of contextual factors when we engage in epistemic evaluation. If some delusions are epistemically innocent, we have cases which put pressure on the trade-off view of epistemic costs and psychological benefits, since it may well turn out that in some cases "safeguarding wellbeing and maintaining epistemic functionality go hand in hand" (Bortolotti, 2015, p. 498). The notion of *epistemic functionality* plays an important role in the application of epistemic innocence. The idea is simply this: the term picks out an agent's ability to function well epistemically—an ability which can be compromised by features of the context in which the cognition in question arises. To say that an agent's epistemic functionality is *restored* is to say that that which was negatively affecting the agent's ability to function well epistemically has been removed or its effect ameliorated. This restoration is thus epistemically beneficial.

There are also clinical implications for understanding delusions in the framework of epistemic innocence. If the case is one in which the delusion meets the conditions on epistemic innocence, then challenging the delusion may be ill-advised since it may well serve an important epistemic function (Bortolotti, 2016, p. 897).

So we have a multitude of theoretical benefits accrued to a framework of epistemic evaluation which has the notion of epistemic innocence in its toolbox.

Though to repeat: my project is not to argue *that* some cognitive states are epistemically innocent (though there will be some such necessary argumentation along the way, since showing that all empiricists are entitled to an outlandish or implausible claim will hardly be of interest). Rather, I am asking, once we settle *that* question: are the theoretical advantages of a positive answer available across the board? The disagreement between one- and two-factor empiricists lies with respect to whether delusion formation involves abnormal belief formation or evaluation. We might understand the debate as one over the epistemic capabilities of subjects with delusions, or indeed over whether these subjects are epistemically irresponsible in a distinctive way. Understood in these terms, a question regarding the applicability of the notion of epistemic innocence—one regarding epistemic evaluation—very naturally arises. Given the different background commitments of one- and two-factor empiricists, are any of them entitled to appeal to the notion of epistemic innocence, and reap the theoretical benefits of it? My claim in this paper is thus a rather surprising one: both one- and two-factor theories of monothematic delusion formation license a judgment of epistemic innocence for some monothematic delusions. Even though these theories diverge markedly on their psychological and epistemological positions on the etiology of delusion, both can appeal to the notion of epistemic innocence.

At first blush, opposing empiricist views suggest rather different evaluations of the epistemic status of monothematic delusions. For example, given that two-factor accounts identify an abnormality in belief formation or evaluation mechanisms, they may well be taken as identifying something which seriously hinders epistemic benefits, in a way that the one-factor account might eschew. If it turns out that despite this major difference between the views, both can appeal to a nuanced characterization of the epistemic status of these cognitions, that would be illuminating. We thus find in this notion something not only extremely important for the reasons outlined above and expounded elsewhere, but something which can unify otherwise opposing positions on delusion formation.

2. Monothematic delusion

Delusions are characterized as "fixed beliefs that are not amenable to change in light of conflicting evidence" (DSM-V, 2013, p. 87). Focus on the epistemic features of delusions has also been the way many philosophers and psychologists have approached their characterizations of them.

Subjects with delusions are said to form their delusions on evidence which does not properly support their content, to maintain them in the face of counter-evidence or counterargument, and delusions may also be incompatible or badly integrated with the subjects' other beliefs (Bortolotti & Broome, 2008, p. 822).

We can categorize an otherwise fairly heterogenous set of beliefs in several ways. One way is to focus on their theme, and distinguish *monothematic* delusions (those concerning a single theme), from *polythematic* delusions (those

concerning multiple themes). Monothematic delusions "can present in isolation in people whose beliefs are otherwise unremarkable" (Coltheart, Langdon, & McKay, 2007, p. 642), and subjects with monothematic delusions do not display more general delusional belief formation. My discussion is limited to the formation of monothematic delusions since it is these delusions that originally motivated some empiricist theories (e.g., Stone & Young, 1997, pp. 329–330). That is not to say that empiricist theories of delusion formation do not have application in the polythematic cases, but focusing on simpler cases is useful for discussion, and indeed these cases are ones which have been the primary focus in the literature. It is also not to say that the class of monothematic delusions is easily delineated; there are cases of patients shifting between delusions (see Gerrans, 2002, p. 50 for a case of a subject shifting between Capgras and Cotard delusion depending on mood). Empiricism, insofar as it focuses on monothematic delusions, is a research program with taxonomical challenges. However, I do not defend any empiricist account here, and so I focus on the simpler cases, since in doing so it is easier to see the relevant features for my project.

Monothematic delusions exhibit the poor epistemic features outlined in my general characterization of delusion, and are also sometimes inconsistent with the subject's behavior. However, though there may be much evidence against the delusional belief, monothematic delusions are ones which are accompanied by anomalous experiences, which have been thought to play an explanatory role in their formation. Such experiences might constitute a source of evidence for the delusion. Monothematic delusions also have serious pragmatic costs: they adversely affect wellbeing in various ways. They might interfere with one's relationships, and one might not be treated as a trustworthy epistemic agent and as such, may be socially sanctioned (Bortolotti, 2015, p. 493).

I assume a doxastic account of monothematic delusions. The case for doxasticism (for delusions in general) in light of arguments against it has been persuasively made (see, for example, Bayne & Pacherie, 2005; Bortolotti, 2010, 2012). Those not persuaded should note that my focus in this paper is on empiricist accounts of delusion formation, proponents of which work from the assumption that monothematic delusions are beliefs. Here, I follow that approach.[1]

3. Epistemic innocence

Epistemic innocence is used to capture the status of a faulty cognition which is both epistemically costly and epistemically beneficial. An analysis of epistemic innocence encourages us to reflect on the relationship between a cognition's psychological and epistemic benefits (Bortolotti, 2015, p. 490). When we accommodate contextual factors in epistemic evaluation, we can resist the trade-off view popular in discussions of the status of faulty cognitions, which has it that psychological benefits come at the expense of epistemic ones, so though it may be psychologically beneficial to have certain cognitions, those psychological benefits

are had by cognitions with significant epistemic costs. The notion of epistemic innocence allows for a more nuanced characterization of this relationship, namely, that in some cases psychological benefits had by epistemically faulty cognitions can lead to significant epistemic benefits.

An epistemically faulty cognition is epistemically innocent if, at a given time, it (i) endows some significant epistemic benefit (*Epistemic Benefit Condition*) onto the subject, which could not be otherwise had, because (ii) alternative, less epistemically faulty cognitions are in some sense *unavailable* to her at that time (*No Alternatives Condition*) (see Bortolotti (2015, 2016), and Sullivan-Bissett (2015) for discussion). In the next two sub-sections, I elucidate these conditions.

3.1. Epistemic Benefit Condition

Meeting the Epistemic Benefit Condition is not to be epistemically *free from faults*. This is a tall order for any cognition, but to claim that a cognition characterized from the outset as *epistemically faulty* is epistemically free from faults would be incoherent. Nor does meeting this condition show that a cognition is epistemically *good* overall. To say this would be to ignore or deny the obvious epistemic costs associated with the cognition under inspection. The Epistemic Benefit Condition does not pick out the result of some weighing procedure: to be epistemically innocent is not to have *more* epistemic benefits than epistemic costs. Epistemic *innocence* does not track epistemic *goodness*, but is used to capture the epistemic standing of a cognition which has a kind of in-between status. Ideally, an agent would not have epistemically faulty cognitions, but human agents have significant limitations that, in certain contexts, lead to such cognitions. The notion of epistemic innocence allows us to pick out those contexts in which such cognitions can play a positive epistemic role.

What counts as an epistemic benefit depends on one's account of epistemic value. I will understand the idea of epistemic benefit in consequentialist terms, whereby the goal of epistemic evaluation is the maximization of epistemic value. The idea is that if the subject did not have the epistemically faulty cognition she would not enjoy such a benefit because alternative cognitions which could confer it are unavailable.

Although I assume a consequentialist framework for the sake of the exposition of ideas, the utility of the notion of epistemic innocence is not restricted to such a framework. As Bortolotti notes, epistemic innocence can also be understood in a deontological or virtue epistemology framework. A deontologist, though not concerned with epistemic benefits, may be interested in whether the faulty cognition were the only one available to the agent at the time, since if epistemically better cognitions were unavailable, this will have consequences for that agent's epistemic responsibility or blameworthiness for adopting the faulty cognition. The evaluation of a cognition as epistemically innocent will speak to the deontologist's interests here, since part of that evaluation is an investigation into whether alternative cognitions were available to the subject (see Section 3.2). A virtue

epistemologist might reframe the notion of epistemic benefit into a claim about which cognitions contribute to the development of epistemic virtues (Bortolotti, 2016, p. 11). A cognition being epistemically innocent will be the kind of thing the virtue epistemologist is interested in capturing, insofar as the epistemic benefits had by a cognition are ones which promote the agent's epistemic virtues.

Finally, in many of the discussions of the Epistemic Benefit Condition in the literature so far, the benefits identified are ones which are had indirectly, as a result of psychological benefits facilitating the epistemic capabilities of the subject. It is this relationship between psychological and epistemic benefits which puts pressure on the trade-off view thereof. My discussion will continue this trend. The epistemic benefits that accrue to a subject forming a monothematic delusion are ones which are mediated by the psychological benefits of doing so. Part of my brief defense of the epistemic innocence of some monothematic delusions will thus suggest that in some contexts these cognitions are psychologically beneficial.

3.2. No Alternatives Condition

The notion of *unavailable* in play in the No Alternatives Condition is apt to give rise to several readings. We can distinguish three ways in which alternative epistemically better cognitions may be unavailable to the subject, which reflect the various limitations subjects might face (see Sullivan-Bissett, 2015, p. 554). The way monothematic delusions meet the No Alternatives Condition will vary depending on whether one adopts a one- or two-factor account of their formation and maintenance. For example, if the subject with a monothematic delusion has a reasoning bias or deficit (two-factor theory), the way in which alternative beliefs are unavailable to her may go via this bias or deficit. If the reasoning of a subject with a monothematic delusion is within the normal range (one-factor theory), the sense in which alternative beliefs are unavailable to her will be different, insofar as it reflects this.

An alternative cognition is *strictly unavailable* if it is inaccessible to the subject in the strongest terms, that is, if it is based on information which is opaque to introspection or otherwise irretrievable. Consider a subject with dementia who suffers from a severe memory impairment. She claims to remember a trip to the beach with her parents that morning. That event did occur, but it occurred sixty years previously. A more accurate cognition is not available to the subject due to the memory impairment she suffers as a result of her dementia (Bortolotti and Sullivan-Bissett, forthcoming). In a case like this, alternative epistemically better cognitions are *strictly unavailable.*

An alternative cognition is *motivationally unavailable* if motivational factors inhibit or make it inaccessible to the subject. Consider a case of wishful self-deception (one in which a subject desires that *p* and, in part because of this desire, comes to believe that *p*). This motivational element of the belief formation process makes less epistemically faulty cognitions unavailable to the subject. Consider someone

strongly motivated to believe that her partner is faithful, despite mounting evidence of his infidelity. An alternative belief (in his infidelity) is motivationally unavailable to her due to her very strong motivation for it to be the case that he is faithful. She thus reinterprets any putative counterevidence in line with this.

Finally, an alternative cognition is *explanatorily unavailable* in cases in which it strikes the subject as implausible enough to not be regarded as a genuine contender. If an alternative cognition is unavailable in this sense, it is either one that the subject happens not to consider (perhaps because of lack of relevant expertise), or one which she rules out on grounds of implausibility or explanatory inadequacy relative to the adopted cognition.[2] Consider an Intelligent Design proponent looking out onto the world and seeing its apparent order. In particular, when she thinks about the complexity of the human eye, she comes to the conclusion that this is the result of the creation of an intelligent designer. An alternative explanation for her evidence is that intention-blind processes of natural selection operating over millions of years of evolution "just so happened" to produce the human eye. This explanation is explanatorily unavailable to her insofar as she would rule it out as implausible, relative to the preferred and adopted cognition.

In many cases of explanatory unavailability, the adopted cognition is the epistemically better cognition, given the subject's evidence. This is because explanatory unavailability of alternative hypotheses often sorts the good from the bad. Unlike in cases of strict or motivational unavailability, where epistemically better cognitions are sifted due to non-truth tracking features of the belief forming processes, in cases of explanatory unavailability the epistemically less good explanations are often the ones that are filtered out. The example I use here is not like this, and we will see later that explanatory unavailability does not always sift out less epistemically good alternative cognitions.

We can distinguish the claim that there is no *accessible alternative* from the claim that there is no *accessible basis for an alternative*. Both kinds of unavailability occur in the case of strict unavailability. It is only strict unavailability which can apply both to the alternative cognition and information which would suggest that the alternative cognition is an epistemically better one. With motivational unavailability, information can only be motivationally *interpreted* (for example, a subject may treat available evidence in a biased way, but such evidence cannot be motivationally unavailable in the sense I have in mind here). Likewise with explanatory unavailability, it is the alternative *cognition* which is unavailable in this sense, not information which would suggest that alternative might be a better one. It is only explanations which can be explanatorily implausible, not information.

Of course, these kinds of unavailability admit of degrees, depending on the nature of the cognitive or motivational limitations and the context in which they occur. For example, if I am *strongly* motivated to believe that my partner is faithful, the degree to which alternative explanations for his behavior are unavailable to me may well be higher than the degree to which those alternatives are unavailable if I have a mild motivation to believe that he is faithful. That there are such

differences in how a case might meet the No Alternatives Condition underlines the importance of contextual factors in epistemic evaluation. And of course, moving to the domain of motivational monothematic delusions, we may find motivational factors together with perceptual abnormalities or reasoning impairments. Relative to these more extreme cases, the non-availability of alternative explanations in the case of self-deception might look less plausible (see Bortolotti, 2015, p. 498 for discussion).

3.3. Why innocence?

I use the term *epistemic innocence* to capture the status of a cognition which is epistemically faulty, but nevertheless has epistemic benefits which are otherwise unavailable. The suggestion is that the evaluation of beliefs should take into account the context in which they emerge, and ask whether alternative beliefs are available to the subject at that time.

I follow Bortolotti (2015, 2016) in understanding *innocence* in terms of *permissibility*. This kind of innocence is used in the legal context in the U.K. and the U.S. An act is *permissible* if it does not constitute a crime because, in the given circumstances, it prevented a greater harm from occurring. To use Bortolotti's example (following Greenawalt, 1986, p. 89): Ann hits Ben which prevents him from detonating a bomb. Though Ann's action is objectionable, since she causes injury to Ben, it is also an action which is beneficial insofar as it prevents a greater harm from occurring. Additionally, the action of hitting Ben may have been the only one available to Ann which had the consequence of preventing the harm that would have been caused by the detonation of the bomb (Bortolotti, 2015, p. 495).

This kind of defense might be usefully appealed to in accounting for cognitions which are epistemically faulty, but also bestow some epistemic benefit by preventing worse epistemic consequences. A notion of innocence in the epistemic domain allows us to pick out contexts in which an epistemically faulty cognition might qualify as a response to an emergency and avoid worse epistemic consequences. The cognition might do this by delivering some epistemic benefit which is not otherwise available (Bortolotti, 2015, p. 497).

4. Empiricisism about monothematic delusion formation

Here, I briefly outline the one- and two-factor approaches to monothematic delusion formation,[3] following John Campbell (2001) in referring to them as *empiricist* approaches. My survey is far from exhaustive, though my claim is that whatever empiricist approach we take, the verdict of epistemic innocence for some monothematic delusions is available. This means that both one- and two-factor theorists, when in the business of epistemic evaluation, can appeal to the notion of epistemic innocence to tell a richer story about the epistemic status of monothematic delusions.

I put aside rationalism and prediction-error accounts. Rationalists deny experience a role in delusion formation, and claim instead that 'delusion is a matter of top-down disturbance in some fundamental beliefs of the subject, which may consequently affect experiences and actions' (Campbell, 2001, p. 89). The epistemic benefits I suggest for monothematic delusions are a result of coming to an explanation of an anomalous experience, or the attendant anxiety relief from having found such an explanation, and the consequential reinstated epistemic functionality. As above, to have one's epistemic functionality *reinstated* or *improved* is just to be enabled to function better in this domain. Thus, an ascription of epistemic innocence to some monothematic delusions in the hands of the rationalist will look very different, since she denies experience a role in the formation of a delusion. It is in part the convergence on experience as a factor in the explanation of monothematic delusion formation which allows for a shared judgment of epistemic innocence for all empiricists, and thus principally excludes rationalism from the discussion. In addition, the case made against rationalism from various quarters is a formidable one (e.g., Bayne & Pacherie, 2004a, 2004b), and so even if it is a theory which could appeal to epistemic innocence, it is one with comparatively little support, and thus any conclusions regarding it will have comparatively little interest.

Prediction-error theories have in the background the claim that perceptual processing involves generating predictions about sensory input based on hypotheses about the world. To minimize the error of these predictions on the basis of comparison between them and sensory input, hypotheses are updated. Delusions are the result of a malfunctioning of this process, whereby erroneous updating occurs and because of continuing faulty signals supporting the updated hypothesis, the delusion persists in the face of counterevidence (Corlett et al., 2007, p. 2396; Corlett, Honey, & Fletcher, 2016, p. 2; Corlett, Taylor, Wang, Fletcher, & Krystal, 2010, pp. 355–357). I exclude prediction-error accounts from the discussion for two reasons.

First, it is unclear where these accounts sit with respect to the number of factors involved in monothematic delusion formation. Some proponents see themselves as single-factor theorists, but at least some of what they say is suggestive of a second factor playing a role. Sarah Fineberg and Philip Corlett understand beliefs 'as akin to stimulus-response habits that become resistant to contradictory evidence through overtraining' (Fineberg & Corlett, 2016, p. 2). In distinguishing themselves from one-factor accounts which posit a single factor (perceptual dysfunction) and two-factor accounts which posit two factors (perceptual dysfunction and belief evaluation deficit), they note that on their account these 'two factors are strongly interrelated' (Fineberg & Corlett, 2016, p. 5). They propose a single impairment in prediction error, occurring in three stages. The first is *delusional mood,* in which "attention is drawn to irrelevant stimuli," the second is *delusion formation*, in which "explanatory insight occurs and flexible processing is disabled," and the third is *explaining things with the delusion*, in which the delusion

becomes habitual, and "enables patients to stay engaged with the environment and exploit its regularities, though the patient may be inflexible and unresponsive to corrective feedback" (Fineberg & Corlett, 2016, p. 4). It looks like the error can characterize either factor, and Fineberg and Corlett themselves note that on their model "top-down and bottom-up processes sculpt one another" (Fineberg & Corlett, 2016, p. 5).

In general terms, prediction error accounts are one-factor accounts insofar as it is the malfunctioning of the prediction error process which is thought to be the source of the error leading to delusional belief. Whether the source of the error occurs in perceptual processes or in later cognitive processes, there is just one kind of factor in play. However, though prediction-error theorists often deny that there is a sharp perception-cognition divide, some nevertheless recognize that in some contexts, prediction error signals will be discounted (e.g., Corlett et al., 2016, p. 4). In light of this, it has been claimed that prediction error theories are compatible with the two-factor approach (Coltheart, 2010, p. 25; Miyazono, Bortolotti, & Broome, 2014). Thus prediction error theorists operate at a different level from the one in the debate between one-factor and two-factor empiricists (for more on this issue see Noordhof and Sullivan-Bissett, manuscript).

Second, prediction-error accounts take their remit to be much wider than the relatively narrow remit of empiricist accounts which only focus on *monothematic* delusions. My focus then is on just one- and two-factor empiricist accounts, since the contrast between these accounts is clearest, which makes for a more interesting conclusion: if these two clearly opposing accounts both license a judgment of epistemic innocence for monothematic delusions, that would be quite a result. Although empiricists may well have hopes that their accounts extend to polythematic cases (e.g., Coltheart, 2010, p. 18, pp. 24–25), their current remit as demonstrated by much of the literature, as well as their original motivation, extends only to the class of monothematic delusions.

All empiricists agree that subjects with monothematic delusions have anomalous experiences which play a role in explaining the formation of the belief. There is disagreement though with respect to the extent to which such experiences can do this explanatory work. One-factor proponents claim that the only clinically significant abnormality afflicting subjects with monothematic delusions is their anomalous experience, there is no further cognitive abnormality to be found in their mechanisms of belief formation or belief evaluation, the psychology thereof is within the normal range. Two-factor accounts locate two abnormalities in subjects with monothematic delusions. The first is the anomalous experience appealed to by one-factor theorists, thought to determine content, the second is some clinically significant cognitive bias or deficit of the belief formation or belief evaluation mechanisms, thought to explain the formation or maintenance of the belief. The precise characterization of the second factor is what distinguishes different two-factor accounts. In the next two sub-sections, I lay out these two accounts.

4.1. One-factor accounts

One-factor accounts claim that a subject forms a delusion on the basis[4] of an anomalous experience, and that the reasoning from that experience is within the normal range; "delusional beliefs are developed in much the same way that normal beliefs are" (Maher, 1988, p. 22).[5] On Brendan Maher's account, the experiences had by subjects with delusions are such as to distort the evidence they have available to them, and so the delusion is not held in the face of obvious counterevidence, rather it is held "because of evidence powerful enough to support it" (Maher, 1974, p. 99).

The kinds of anomalous experiences subjects with monothematic delusions might have include "a discrepancy between intention and action" (e.g., delusions of alien control in which a subject's bodily movement is not intended by her), as well as "hallucinations, substantial impairments of perceptual processes such as depth perception, the constancies of size, color, shape, etc., and the incapacity to exclude distracting input of various kinds" (Maher, 2003; p. 18). For example, a subject with perceptual delusional bicephaly (the delusion that one has a second head) may hallucinate a second head on her shoulder. Not all anomalous experiences are hallucinatory in this way though (not all present objects and properties as in the world when they are not). In Capgras delusion for example (the delusion that a loved one has been replaced by an imposter), the anomalous experience is one of absence—subjects have reduced affective response to familiar faces due to ventromedial damage (Davies, Coltheart, Langdon, & Breen, 2001, p. 144).

Maher claims that a single experiential factor can also explain why delusions are maintained. To demonstrate that tenacity and seeming insensitivity to counterevidence is not something which only characterizes delusional belief, Maher compares the retention of a delusional explanation for an anomalous experience to the retention of scientific theories:

> As in science, a coherent theory is only overthrown by a better theory and the chances that this can be done successfully by a clinician are reduced when the patient has found a generally satisfactory theory of his own. (Maher, 1974, p. 107)

A delusional belief then is like many others in its being maintained until a better hypothesis is available which explains the data. We do not need to appeal to a second factor to explain why delusional beliefs are maintained, rather, their being so aligns subjects with delusions with non-delusional subjects, who are also resistant to give up their theories in the face of counterevidence (Maher, 1988, p. 22).

4.2. Two-factor accounts

For those who think that anomalous experience plays a role in the explanation of delusion formation, two-factor accounts represent the current orthodoxy. What characterizes all such accounts is the claim that although anomalous experiences are involved in the formation of a monothematic delusion, we need to appeal

to something abnormal at the level of belief formation or evaluation for a full explanation. Here, I very briefly outline some versions of the two-factor approach which I will refer back to later when discussing the two-factor approach and epistemic innocence. I do not consider every version of the approach here. But the arguments given in the next few sections can, with appropriate amendments, have application when one substitutes one of the two-factor accounts I do discuss with one that I do not.

Philippa Garety and Daniel Freeman claim that "there is growing evidence of reasoning and attributional biases in people with delusions which suggests they may display systematic differences in cognitive processes from those in the general population" (Garety & Freeman, 1999, p. 116). They claim to have found "strong support" for there being a reasoning bias which they describe as a data-gathering bias. This bias is understood as a tendency to gather less evidence than non-delusional subjects, such that subjects with delusions "jump to conclusions" (Garety & Freeman, 1999, p. 147). (See for example, Huq, Garety, & Hemsley, 1988; Garety, Hemsley, & Wessely, 1991; cf. Ross, McKay, Coltheart, & Langdon, 2015.)

Tony Stone and Andrew W. Young identify the second factor as the tendency to privilege observational data over minimizing adjustments to one's beliefs (1997, pp. 349–350). They follow Nicholas Humphrey (1995, p. 30) in understanding conservatism about belief fixation in terms of three principles: *frugality* ('Try to explain things in terms of previously established principles, rather than by introducing new ones on a purely ad hoc basis'); *elegance*, ('Try to explain things with a minimum of special pleading'); and *probability* ('Prefer the probable to the improbable. Try to explain things in whatever way seems least unlikely'). Stone and Young claim that in subjects with delusions, the balance between observational adequacy and conservatism "goes too far" toward privileging observational data (Stone & Young, 1997, p. 349).

Martin Davies and colleagues claim that we need to explain why a delusion is *maintained,* as well as why it is *formed* initially, given its "implausibility and its inconsistency with everything else that the patient knows" (Davies et al., 2001, p. 151). They understand the second factor in terms of a failure to inhibit a *pre-potent doxastic response.* Though the subject's belief *formation* is normal, there is a deficit in the belief *evaluation* system. We usually believe what we perceive, but normal subjects can "suspend their unreflective acceptance of veridicality and make a more detached and critical assessment of the credentials of their perceptual experiences" (Davies et al., 2001, p. 153). Davies and colleagues suggest that:

> hypotheses that were generated from the patient's own perceptual experience would be resistant to being critically assessed and recognized as implausible, but hypotheses generated by someone else would be assessed in the normal way. (Davies et al., 2001, p. 153)

This second factor is understood by Davies and colleagues as "the loss of the ability to reject a candidate for belief on the grounds of its implausibility and its inconsistency with everything else that the patient knows" (Davies et al., 2001, p. 154).

Max Coltheart furthered this view, suggesting that the second factor is "a (very poorly specified) defect of a belief evaluation system," located in the right hemisphere (Coltheart, 2005; p. 154). In all cases of monothematic delusion, there are two neuropsychological impairments. First, an impairment which differs from one patient to another and "presents the patient with new (and false) data" which the delusional belief would explain, if it were true (Coltheart, 2005, p. 154). Second is an impairment in belief evaluation, the same in all subjects with monothematic delusions, which "prevents the patient from rejecting the newly formed belief even though there is much evidence against it" (Coltheart, 2005, p. 154).

Some have distinguished two kinds of two-factor theory: competence vs. performance theories (see, for example, Gerrans, 2001, pp. 166–171; Bayne & Pacherie, 2004a, p. 6). Philip Gerrans agrees with Maher that delusional beliefs are "justified by the application of unimpaired procedural rationality to an anomalous experience" (Gerrans, 2001, p. 162), and that the phenomenology of such experiences is "intense, disturbing, and intractable" (Gerrans, 2001, p. 169). However, he agrees with two-factor theorists like Haydn Ellis (1998) that the delusion is nevertheless irrational and abnormal in some sense. Gerrans reconciles these thoughts by distinguishing between *procedural* and *pragmatic* irrationality. He claims that subjects with delusions understand the requirements placed on them by procedural rationality (they recognize that others might regard their delusion as inconsistent with their other beliefs), and are nevertheless "prevented from adopting the normal conclusion" (Gerrans, 2001, p. 165). The idea is that subjects with delusions differ from the non-delusional population with respect to pragmatic rationality, which is the faculty by which normal subjects select between hypotheses which are equally consistent with the prima facie evidence.[6] Pragmatic rationality is the faculty employed to answer questions such as 'What counts as good evidence? … How are initial probabilities assigned?' (Gerrans, 2001, p. 162). In some contexts a reasoning *performance* can be such that it does not reflect a subject's reasoning *competence*: for example, JK had Cotard delusion (claiming that she was dead). However, she was asked if she could feel her heart beating, and claimed that a beating heart was no sign of her being alive, given that she was both dead and had a beating heart. She also recognized that other subjects would find this hard to believe (Young & Leafhead, 1996, p. 158). Gerrans's diagnosis of this case is that JK's competence is intact but her performance is in error (Gerrans, 2001, pp. 167–171).[7]

It is unclear whether Gerrans's model is best characterized as involving one or two factors, that is, whether the pragmatic irrationality is an extra factor, or whether it is bound up with the anomalous experience. Gerrans says something which is closer to the latter option. He says of the subject with delusions who displays competence with respect to procedural irrationality that she:

> [has] a performance deficit possibly based in the cause of her anomalous experience, which, as one-stage theorists often emphasise, is both extremely distressing and cognitively intractable. (Gerrans, 2001, p. 170)

I read Gerrans as a two-factor theorist, since he identifies as "*the difference* between delusional and normal subjects" their pragmatic rationality, thus explicitly drawing a line between subjects with delusions and subjects without with respect to reasoning (Gerrans, 2001, p. 162, emphasis added).[8]

Finally, drawing lessons both from Davies and colleagues' claim about pre-potent doxastic response and Gerrans's claim about performance, Jakob Hohwy and Raben Rosenberg suggest that in delusion there is a failure to inhibit the pre-potent doxastic response, and this ought to be understood as a localized performance failure of the relevant cognitive mechanism *in specific circumstances* (Hohwy & Rosenberg, 2005, p. 143). This account is said to avoid the unhappy prediction that subjects with delusions should form all sorts of beliefs in light of unusual experiences due to a more global deficit of this kind (Hohwy & Rosenberg, 2005, p. 142).

We have seen some ways of characterizing the second factor putatively involved in monothematic delusion formation: in terms of a data-gathering bias, in terms of a bias toward observational adequacy over conservatism, or as an inability to reject candidates for belief because they are implausible or inconsistent with the subject's other beliefs. One way of understanding this inability is offered by Coltheart, who claims that there is a deficit in belief evaluation which prevents the delusion being rejected. Finally, moving from the second factor as a competence error to it as a performance error, Gerrans claims that in some contexts subjects with delusions suffer from a performance error in their reasoning, even though their competence is generally intact, and Hohwy and Rosenberg, working in the framework of performance error, understand the failure to inhibit a pre-potent doxastic response in these terms.

5. The epistemic innocence of monothematic delusions and empiricism

I now move to considering whether a judgment of epistemic innocence for monothematic delusions is something upon which all empiricists about their formation can agree. I will argue that it is.

5.1. Epistemic benefits of monothematic delusions

Monothematic delusions fail to meet many epistemic standards, as we saw in the characterization I gave earlier. So as not to go over this well-trodden ground any further, I move now to a discussion of the epistemic *benefits* of monothematic delusions, some of which are identified by Bortolotti (2015). These benefits are mediated by psychological benefits and support the epistemic functionality of the subject, which might otherwise remain compromised. They are also related to the anomalous experiences subjects with monothematic delusions have. In this context, the term *epistemic functionality* picks out an agent's ability to function well epistemically, which, depending on one's preferred framework of epistemic evaluation, might be the ability to acquire true beliefs or knowledge, to exercise

one's intellectual virtues, and so on. To have one's epistemic functionality improved or restored is thus epistemically beneficial.

I reiterate that my aim here is not to make a definitive case for the epistemic innocence of some monothematic delusions (some of that work has been done, see Bortolotti 2015), but rather to argue that if epistemic innocence does have applicability in this context, that is a claim licensed by opposing empiricist theories of the formation of these cognitions.

5.1.1. Apparent gap-filling and epistemic functionality

Anomalous experiences can create in a subject "puzzlement, anxiety, and a search for an explanation" (Maher, 2006, p. 181). In the face of an experience which the subject cannot explain, the formation of a delusion might fill an explanatory gap, providing an explanation of an as yet unexplained phenomenon (I use the term 'apparent' since one might think that *filling an explanatory gap* is a success term). As Coltheart puts the point in his discussion of explanationist accounts of monothematic delusions:

> a state of the world has arisen for which the patient has to find an explanation, and the delusional belief provides such an explanation: that is, if the delusional belief were true, then it would follow that the world would be the way it now seems to be to the patient. In this sense, the belief explains why the world is as it now seems to be. (Coltheart, 2005, p. 153)

The subject then no longer faces an explanatory gap caused by anomalous experiences. We might naturally think that gap-filling is epistemically beneficial only when it tends toward truth, and so the mere *apparent* gap-filling function performed by the delusion is not epistemically beneficial. There are two routes to the claim of epistemic benefit here. First, it might be that with the delusional belief in place, other beliefs of the subject are made coherent. This would be an epistemic benefit available only to those who thought mere coherence was an epistemic end in itself. Second, the epistemic advantage accrued via apparent explanatory gap filling is one mediated by psychological benefits: apparent gap-filling might have downstream positive epistemic consequences, mediated by the psychological benefits thereof.

In her discussion of delusions in schizophrenia, Bortolotti claims that a delusion "may support epistemic functionality through engendering a new attitude towards experience. The agent … no longer finds her experience puzzling, but feels that it is in her power to understand it and that it is important to come to such an understanding" (Bortolotti, 2016, p. 18). This point is also applicable to monothematic delusions, since they too are ones associated with anomalous experiences which may be puzzling to the subject. On the broadly consequentialist view of epistemic value I adopt here, we can understand the formation of a monothematic delusion as supporting an agent's epistemic functionality (which is compromised by the anomalous experience she has). When the explanatory gap left by these experiences is considered closed, epistemic functionality is restored.

Support for the claim that the perceived filling of an explanatory gap is psychologically beneficial might be drawn from Bergstein, Weizman, & Solomon, 2008 study on the Sense of Coherence (SOC) in subjects with delusions. SOC picks out a:

> global orientation that expresses the extent to which one has a pervasive, enduring though dynamic, feeling of confidence that (1) the stimuli deriving from one's internal and external environments are structured, predictable, and explicable; (2) the resources are available to one to meet the demands posed by these stimuli; and (3) these demands are challenges, worthy of investment and engagement. (Antonovsky, 1987, p. 19)

Bergstein and colleagues found that the average SOC score for patients at the acute stage of experiencing delusions was similar to the average SOC score of the non-delusional population. A six month follow-up revealed that those subjects in remission, had *lower* SOC scores than those subjects who remained delusional, a finding Bergstein and colleagues note as one which "contradicted common assumptions about breakdown of coherence of experience during chronic psychotic states" (2008, p. 291). Furthermore, Bergstein and colleagues refer to work which shows that depression is associated with remission from delusional states. They suggest that this might be an unsurprising finding "if remission from delusional states is accompanied by a reduction in the level of meaning and purpose … with a resultant decrease in psychological well-being" (2008, p. 289).

If SOC is increased *as a result of having a delusion,* we might speculate on some downstream positive epistemic consequences of this. If I am a subject who is highly confident that "the stimuli deriving from [my] internal and external environments are structured, predictable, and explicable" and that "the resources are available to [me] to meet the demands posed by these stimuli," this may increase my confidence in forming beliefs and gaining knowledge on the basis of such stimuli. If high SOC is indirectly epistemically beneficial, and SOC is not significantly different in subjects with delusions than in subjects without, *and higher* than subjects in remission, this suggests that delusions are not faulty with respect to the SOC of subjects who have them. Rather, those subjects in remission have *reduced* SOC, suggesting that with respect to SOC, the subject not having the cognition (remission) would be epistemically less advantaged than one having it, or not significantly different (normal population).[9] In forming a delusional belief then, subjects enjoy some psychological advantages insofar as these cognitions seem to fill an explanatory gap, and these psychological advantages can lead to positive epistemic consequences.

5.1.2. Anxiety relief and epistemic functionality

The restoration of epistemic functionality may also result from anxiety relief. Before coming to a belief regarding their anomalous experiences, it should be kept in mind that subjects in these contexts may have been in a state of considerable distress. For example, in the context of discussion of the prodromal stage of psychosis and the delusional atmosphere thereof, Karl Jaspers suggests that patients "suffer terribly under it and to reach some definite idea at last is like being relieved

of some enormous burden" (1963, p. 98). If monothematic delusions are accompanied by highly anomalous experiences, similar things apply in this context. Maher suggests that coherent explanations of anomalous experiences are often accompanied by a "strong feeling of personal relief," as well as the excitement a scientist might get from intellectual insight (Maher, 1974, p. 104).

Negative emotions might be managed by the formation of a delusion. For example, Peter Butler argues that "the emergence of erotically themed delusions following trauma may represent an active attempt to regain intrapsychic coherence and to confer meaning on otherwise catastrophic loss or emptiness" (Butler, 2000, p. 85). In his case study of B.X., a subject with Reverse Othello syndrome, believing that he was happily married when his partner had left him, Butler reports that "B.X.'s fantasy system functioned to protect him from the consequences of massive narcissistic injury and attendant depressive overwhelm" (Butler, 2000, p. 88). David Raskin and Kathleen Sullivan (1974) draw on cases studies of erotomania and suggest that the delusion "served the purpose of warding off depression and loneliness" (p. 1034), as well as "fill[ing] a vacuum" in the patients' lives (p. 1035).

Could the anxiety felt before coming to the delusional explanation for anomalous experience be such as to adversely affect wellbeing and compromise the gaining of knowledge? Bortolotti claims that:

> Consequences of stress and anxiety include lack of concentration, irritability, social isolation, and emotional disturbances. These in turn negatively affect socialisation, making interaction with other people less frequent and less conducive to useful feedback on existing beliefs, and to the fruitful exchange of relevant information. Due to reduced socialisation and engagement, the acquisition and retention of knowledge is compromised and intellectual virtues are not exercised. (Bortolotti, 2015, p. 496)

We see here that the psychological benefits of the adopted cognition may translate into epistemic ones. Subjects with monothematic delusions are often distressed prior to the adoption of the belief which explains their experience. The formation of the delusion may restore epistemic functionality by providing relief from anxiety as well as intellectual satisfaction in explaining their experience (an "aha moment" [Conrad, 1958, cited in Mishara, 2010, p. 10]).

In the case of motivated delusions, negative emotions might be managed which, in the absence of the delusion, could lead to depression or low self-esteem. The formation of a delusion of this kind might be "restitutive, ameliorating anxieties by altering the construction of reality" (Lansky, 1977, p. 21). In turn, we might expect that insofar as monothematic delusions perform this function, the agent's epistemic functionality will be thereby supported.

I should note the limits of the epistemic benefits attained via relief from anxiety. Though the formation of a delusion may temporarily relieve anxiety caused by anomalous experiences, the maintaining of a delusion often increases anxiety (Bortolotti, 2015, p. 496). There might be relief from anxiety at the prodromal stage, when the subject is puzzling over how she ought to interpret the experience she is having, but once the delusion is formed, a new kind of anxiety is in play,

namely, one linked to the content of the delusion. A subject with Capgras delusion may be anxious about the whereabouts of her husband, a subject with Cotard delusion may be anxious about being dead, and so on. Delusions with negative contents "are correlated with higher depression and lower self-esteem" (Smith et al., 2006; cited in Bortolotti, 2015, p. 496). However, my interest is in whether monothematic delusions, at the very least, *start* as epistemically innocent, and whether this judgment can be licensed by different empiricist accounts of their formation. So this worry is not going to count against delusions *being* epistemically innocent, but only against their *staying* epistemically innocent. Bortolotti notes a similar interest: her discussion takes place in the context of delusional beliefs being adopted "at a time when access to the truth is already compromised by the effects of trauma or previous adversities, and it would be further compromised unless negative emotions were effectively managed" (Bortolotti, 2015, p. 495). It is in contexts such as this we that see monothematic delusions can play a positive epistemic role.

We have seen the indirect epistemic benefits which monothematic delusions might have: the apparent filling of an explanatory gap left by an anomalous experience by providing an explanation of an as yet unexplained phenomenon, and the relief from anxiety as a result of forming the belief. Both of these features of monothematic delusion formation may restore epistemic functionality, that is, the ability to function well epistemically, which may lead to true beliefs and knowledge.

Of course, in some cases the epistemic consequences of resuming engagement with the world enabled by forming the delusion may well be very bad. For example, if the delusion encourages disorganized behavior or if it acts as a skewed framing heuristic for the subject's new beliefs, the formation of the delusion may in fact be epistemically disastrous (see Bortolotti, 2016, p. 892 for discussion of this possibility in the context of delusions in schizophrenia). Given their relative circumscription, monothematic delusions are perhaps less likely to have these kinds of negative epistemic consequences. But in any case, where we find such consequences we have a good case for thinking that the Epistemic Benefit Condition is not met. Epistemic innocence is a feature sensitive to context, and of course there will be some contexts in which it does not apply.

In some cases then, without the delusion (and given the unavailability of alternative hypotheses as I will argue below), we would have something like the following situation: the subject has an anomalous experience which she cannot explain, so there is an explanatory gap (epistemic cost), as well as the downstream epistemic costs associated with anxiety in the prodromal stage. With respect to this situation, the delusion can be seen to confer epistemic benefits, and this is the case whether one is a one- or two-factor theorist.

Might a subject be epistemically better off *withholding* belief on the matter of her experience? Perhaps, but it is no part of the case for epistemic innocence that the agents involved are taking the *ideal* epistemic route, or the one which would

bestow the most epistemic benefits. We might think that when *p* is true, the best thing to do is believe that *p,* the next best thing to do is to withhold belief that *p,* and the worst thing to do is to believe that not-*p* (McHugh, 2012). In the case of Capgras delusion, for example, the best thing epistemically would be to form the belief that one is suffering from a lack of affect to familiar faces brought on by ventromedial damage. The next best thing would be to withhold belief, and the worst thing would be to believe that one is *not* suffering from illness. However, it is no part of the claim of epistemic innocence that subjects forming monothematic delusions are in line with the prescription of epistemic norms, the claim is only that the epistemic route taken is one which has epistemic benefits mediated by psychological ones.

I have sought only to make a plausible enough case for some monothematic delusions meeting the Epistemic Benefit Condition such that my conclusion is an interesting and important one. More needs to be said to make that case definitive. The question now is whether these benefits are otherwise obtainable; that is, whether alternative beliefs are available to the subject with delusions.

5.2. No alternatives to monothematic delusions

As we have seen, the difference in the way monothematic delusions warrant an appraisal of epistemic innocence depending on one's preferred empiricist account does not manifest in the application of the Epistemic Benefit Condition. However, the way monothematic delusions meet the No Alternatives Condition will differ depending on whether one adopts a one- or two-factor approach, which demonstrates the importance of a project like the one undertaken here, one that does not investigate the epistemic innocence of some cognition independently of theories of the formation of such a cognition. An account of why monothematic delusions meet the No Alternatives Condition cannot help but appeal to features of their formation which will be inescapably theory-laden. Further, the ways of challenging a delusion might be informed by reflection on how the delusion meets the No Alternatives Condition, since it is this condition which makes the delusion the only way to attain the epistemic benefits outlined above. By eliminating the contextual features which contribute to the meeting of the No Alternatives Condition, we can learn how to treat the delusion, specifically, to replace it with something with the same epistemic benefits but fewer psychological and epistemic costs.

Next I consider whether monothematic delusions meet the No Alternatives Condition from the standpoint of the one- and two-factor theorist. I will conclude that both accounts license a judgment that they do, but in different ways.

5.2.1. One-factor empiricism and the No Alternatives Condition

One-factor accounts have the resources to support the claim that alternative beliefs are not available to subjects with monothematic delusions. We saw earlier that anomalous experiences are taken by Maher to distort the evidence available to

the subject, a delusion is developed via "evidence powerful enough to support it" (Maher, 1974, p. 99), and that such experiences cannot be "reasoned away." This is not just to say that argument cannot make the experience cease, but rather that such an experience has epistemic import which is not easily undercut by claims about the experience's veridicality.

The epistemic import given by these experiences might go some way towards explaining why subjects with monothematic delusions seem to prefer their hypotheses to naturalistic ones: their experiences are such that they are better explained by the delusional theory. This means that "asking patients to prefer a naturalistic theory to their own" would be "tantamount to asking them to trust the evidence of other people's sense in preference to their own" (Maher, 1988; p. 25). This is something which is "not impossible," but it is also "not readily done by most people" (Maher, 1988, p. 25). The notion of *explanatory unavailability* is in play here, at least for subjects not aware of naturalistic explanations, or for those who are aware that there might be other explanations of their experiences, but nevertheless reject them.

Recall that some two-factor theorists see the explanatory burden as one of explaining delusion *retention,* rather than delusion *formation.* One-factor accounts are claimed to have the resources to meet this explanatory burden, and in their doing so we can identify another way in which alternative beliefs are unavailable. In Maher's explanation of why subjects maintain their delusions in the face of putative counterevidence, something like motivational unavailability is in play. Maher claims that the explanation hit on by a subject with delusions may provide "enough relief from anxiety that it becomes difficult for the individual to abandon it and return to the initial state of confusion and distress. Thus, strong resistance develops to counter-evidence that could rebut the delusional belief" (Maher, 2006, p. 182). We saw in my discussion of the Epistemic Benefit Condition that monothematic delusions can be indirectly epistemically beneficial insofar as anxiety relief can facilitate epistemic functionality. We see here that some one-factor theorists identify this anxiety relief as something subjects may well be loathe to give up, which suggests that alternative cognitions could not support the subject's epistemic functionality in the same way, or to the same extent, as the adopted delusion. Further:

> With the passage of time, the social costs and consequences of major decisions made under the influence of the delusion may create a situation in which it is very difficult for the patient to re-examine the belief and publicly reject it. A patient who has lost employment, been divorced, been alienated from his friends, and so forth has incurred losses that could be justified only if the belief were true. This makes it increasingly difficult for him or her to abandon the belief in spite of mounting evidence of its falsity. (Maher, 2006, p. 182)

Here again we see that alternative cognitions may be motivationally unavailable to the subject, given the costs she has endured as a result of holding the delusion—such costs are only worth bearing if the delusion is true. Thus resistance develops,

and alternative, epistemically less faulty cognitions are unavailable. When it comes to the option of *withholding belief,* we might appeal to motivational unavailability to explain why some subjects do not, indeed *cannot,* simply withhold belief on the matter. Cases in which the anomalous experience is distressing and confusing are ones in which simply withholding belief on the matter of its cause would only serve to keep the anxiety in play.

The one-factor account of monothematic delusion formation licenses the claim that these beliefs, at least sometimes, meet the No Alternatives Condition on epistemic innocence. This is either because alternatives are explanatorily unavailable (since they are considered to be implausible or explanatorily inadequate explanations of anomalous experiences) or because they are motivationally unavailable (due to the anxiety relief given by the delusion, and also the social costs encountered by having it which would only be tolerable if the delusion were true).

5.2.2. Two-factor empiricism and the No Alternatives Condition

Two-factor accounts also license the judgment that monothematic delusions meet the No Alternatives Condition. Here, I consider the versions of the two-factor account I outlined earlier and note again that the arguments here, with suitable amendments, can generate the same conclusions for other two-factor accounts.

I turn first to versions of the two-factor approach which identify the second factor as a data-gathering bias. This might be understood as a tendency to "jump to conclusions" (Garety & Freeman, 1999, p. 147), or the tendency to privilege observational data rather than adjust one's beliefs (Stone & Young, 1997, pp. 349–350). Matthew Broome suggests that cognitive biases such as those identified as the second factor in two-factor accounts may do two things. First, they might support the generation of an anomalous experience, and second, they might increase the chances of a "morbid explanation" of such an experience. A bias …

> … would likely act in limiting the amount of data gathered to support an explanation, and thus end the search for meaning prior to potentially falsifying information being considered. (Broome, 2004, p. 37)

The biases identified by two-factor theorists would do as Broome supposes, that is, prevent the consideration of alternative hypotheses. Such alternatives would, depending on the details of how the bias operates, be unavailable to the subject with delusions.

Consider Davies and colleagues' view that the second factor is one which explains the retention of a delusional belief; they propose it is a failure to inhibit a pre-potent doxastic response. In developing this approach, Coltheart claimed that there are two neuropsychological impairments involved in monothematic delusion formation and retention. The first is involved in the experiential factor, and the second is one which "prevents the patient from rejecting a newly formed belief event though there is much evidence against it" (2005, p. 154). If the subject has a deficit such that she is unable to reject other explanations of her

experiences, then we see that alternative cognitions, rather straightforwardly, are unavailable to her.

We can also look to Coltheart, Menzies, and Sutton's (2010) work on abductive inference and delusional belief, in which they argue that subjects with delusions form beliefs in line with a Bayesian model of abductive inference, according to which "one hypothesis H_1 explains observations O better than another hypothesis H_2 just in case $P(O \mid H_1) > P(O \mid H_2)$" (2010, p. 276).[10] The idea is that a subject's belief, say, that her neighbors are out to get her, is a better explanation of what she observes than an alternative belief if and only if what she observes is more probable under the hypothesis that her neighbors are out to get her than under the alternative, non-delusional hypothesis (Coltheart et al., 2010, pp. 271–272).

Coltheart and colleagues take it that the delusional hypothesis *is* more probable than the non-delusional hypothesis, given the observed phenomena. Considering a case of the Capgras delusion, the two hypotheses in play are the *stranger hypothesis* (the woman who looks like my wife is not my wife) and the *wife hypothesis* (the woman who looks like my wife is my wife):

> The observed data are clearly much more likely under the stranger hypothesis than under the wife hypothesis. It would be highly improbable for the subject to have the low autonomic response if the person really was his wife, but very probable indeed if the person were a stranger. (Coltheart et al., 2010, p. 277)

Coltheart and colleagues suggest that the reason the subject does not reject the delusional hypothesis once the disconfirming data starts to come in is because he seems to be …

> … ignoring or disregarding any new evidence that cannot be explained by the stranger hypothesis. It is as though he is so convinced of the truth of the stranger hypothesis by its explanatory power that his conviction makes him either disregard or reject all evidence that is inconsistent with the hypothesis, or at least cannot be explained by the hypothesis. (Coltheart et al., 2010, pp. 279–280)

Coltheart and colleagues' work here suggests that non-delusional alternative hypotheses are unavailable to delusional subjects insofar as they are *explanatorily* unavailable. The delusional hypothesis just does a much better job of explaining the observed data, if what we mean by "doing a better job" is cashed out in terms of Bayesian probabilities. Care is needed here, lest we make explanatory unavailability too cheap (see note 2). If Coltheart and colleagues are right that monothematic delusions are maintained in this Bayesian rational way, there will be much heterogeneity with respect to whether or not we have explanatory unavailability in play. It is not enough that alternative hypotheses strike the subject merely as less plausible than the adopted cognition, something more is required, specifically viewing alternative hypotheses as *implausible* or with incredulity. Here, again I underline the point that context matters for application of epistemic innocence.

Let us look next to Gerrans's account. We saw in my exposition earlier that his claim is that subjects with delusions understand the requirements placed on them by procedural rationality but are nevertheless "*prevented* from adopting the

normal conclusion" (Gerrans, 2001; p. 165, emphasis added). In his discussion of cases from the literature in which subjects display intact procedural rationality he comments that in such cases "subjects are aware that their conclusion is not one that would be reached by a normal reasoner yet are *blocked* from reaching the normal conclusions themselves" (2001, p. 168, emphasis added). He also claims that, regardless of how extreme an anomalous experience undergone by a subject with a delusion might be, rationality requires that she draw the non-delusional conclusion, and this is something subjects seem to be aware of. We need to appeal to a deficit in pragmatic rationality to explain those cases where "the ability to make this type of judgment is *destroyed*" (Gerrans, 2001, p. 168, emphasis added).

The language Gerrans uses ('prevented', 'blocked', 'destroyed') suggests that cognitions are *strictly* unavailable to the subject, but they could equally be understood in the weaker sense of explanatory unavailability. After all, Gerrans's discussion is of inferential abilities, and the non-delusional conclusion might strike the subject as explanatorily poor, given her defective pragmatic rationality. Recall that such a faculty is one which helps the subject decide between competing hypotheses which are equally consistent with the evidence.

Finally, we find in Hohwy and Rosenberg's account a very natural home for the thought that perceptual impairments of anomalous experiences may well render some information inaccessible. On their view, subjects with delusions fail to inhibit a pre-potent doxastic response and take their anomalous experiences as veridical. They do so because these experiences are such that available reality testing mechanisms produce the same results, and further reality testing mechanisms are not available (Hohwy & Rosenberg, 2005, p. 158).

Let us turn briefly again to the option of withholding belief. I have said that this option is unavailable to at least some subjects forming monothematic delusions, and perhaps in the context of two-factor theories of their formation it is easy to see why. If subjects with monothematic delusions *jump to conclusions,* their doing so may "end the search for meaning" (Broome, 2004, p. 37) rather than withholding belief on the matter. Or, if the second factor is understood as a failure to inhibit a pre-potent doxastic response, the subject is simply *unable* to prevent herself from forming the relevant belief—withholding is simply not an option. Finally, if these subjects suffer from a deficit in pragmatic rationality, they are unable to properly select between competing hypotheses and indeed to properly determine whether one should withhold belief since further evidence should be sought. All of these second factors support the claim that the epistemic route of withholding belief is strictly unavailable to subjects with monothematic delusions.

We have seen that monothematic delusions meet the No Alternatives Condition on epistemic innocence if one adopts a two-factor account of their formation. Though my survey here is far from exhaustive, I am confident that other two-factor accounts will license the same judgment, given that they are set up to explain—via a clinically significant reasoning bias or deficit—precisely why a delusion is adopted rather than an epistemically better belief. As such, whatever that second

factor is, it will do the work of showing that monothematic delusions meet the No Alternatives Condition.

6. Conclusions

I have argued that a judgment of epistemic innocence to some monothematic delusions is one licensed by both one- and two-factor empiricist accounts of their formation. This is an important result since it means all parties can give a richer epistemic evaluation of monothematic delusions, and it also demonstrates the utility of the notion of epistemic innocence, even in the context of opposing views on delusion formation, views which may well be taken to diverge markedly on judgments of epistemic status. There may well be residual worries regarding the applicability of epistemic innocence to monothematic delusions, and indeed worries about the coherence or usefulness of the notion itself. My project here is not to answer these worries. Rather, it is to demonstrate that the judgment of epistemic innocence in this context is one which is compatible with both one- and two-factor empiricist approaches to monothematic delusion formation. I grant that there is surely more to be said about epistemic innocence and monothematic delusion formation. Here, though, I only seek to show that empiricists of various stripes can and should be part of that discussion.

Notes

1. As Bortolotti has noted, the application of epistemic innocence is not restricted to delusions understood in a doxastic framework since even non-doxasticists take delusions to be *belief-like* and thus subject to epistemic evaluation (Bortolotti, 2016, p. 880, note 1).
2. Explanatory unavailability admits of degrees, but it is important not to use this notion in too weak a sense. For example, we ought not say that an alternative explanation is unavailable in this sense if it is merely considered less likely or less plausible than the preferred explanation. Going this weak opens up the door for the worry that alternative explanatory hypotheses are never available when we form beliefs. I am grateful to an anonymous reviewer for seeking clarity on this.
3. Although I remain neutral here on the theoretical success of these accounts, I have defended a one-factor theory at length elsewhere (Noordhof and Sullivan-Bissett, manuscript; Sullivan-Bissett, manuscript).
4. I leave out discussion of exactly what "on the basis of" means here. What we say about that will depend on whether one adopts an endorsement or explanationist account of the relationship between experience and the content of a delusional belief (see Bayne & Pacherie, 2004a, pp. 2–3). For simplicity, I speak in explanationist terms here of delusions as explanatory hypotheses for the experience. I think that endorsement accounts are equally entitled to the idea of epistemic innocence. Indeed, it might even be that the stories are very similar, even down to the kinds of unavailability of alternative hypotheses in play. Endorsing the content of an experience in one's belief is consistent with alternative cognitions being explanatorily unavailable. It is just that in this framework those alternative cognitions are ones which explain why the subject

sees his mother as an imposter, as opposed to lacking some affective response when looking at his mother.

5. Gerrans was a temporary defender of a one-factor position according to which monothematic delusions are "rationalizations of anomalous experiences via reasoning strategies that are not, in themselves, abnormal" (Gerrans, 2002, p. 47).
6. Gerrans has since abandoned this position, and develops a new account elsewhere (2014). His *pragmatic rationality* looks a lot like what José Bermúdez calls *epistemic rationality*, concerned with how beliefs relate to evidence and how they ought to change in response to it (Bermúdez, 2001, p. 468). Similarly to Gerrans, Bermúdez takes an impairment in this domain to be partly explanatory of delusion formation (in Bermúdez's case, for delusions arising in the context of schizophrenia) (Bermúdez, 2001, p. 470).
7. See Noordhof and Sullivan-Bissett (manuscript) for an argument against this interpretation of the case.
8. For the purposes of this paper, it does not matter whether Gerrans is best characterized as a two-factor theorist. If the performance deficit in pragmatic rationality Gerrans refers to is not something over and above the anomalous experience, this does not matter for my argument, since I will argue that whether one is a one- or two-factor theorist, some monothematic delusions are epistemically innocent.
9. What would be useful is work investigating the level of SOC in subjects post-experience, pre-delusion, since if this were lower than it is in subjects with monothematic delusions, this too might indicate that such cognitions have epistemic benefits.
10. My appeal to Coltheart and colleagues' claim here should not be taken to indicate agreement with it. See Ryan McKay for an argument that delusion formation is not Bayesian rational in this way (McKay, 2012, pp. 338–341).

Acknowledgments

I am grateful to audiences at the Mind and Reason research group at the University of York and the Epistemology Seminar at the University of Edinburgh. Thank you to Lisa Bortolotti, Kengo Miyazono, Kathy Puddifoot, and two referees for this journal for very helpful comments on earlier versions of this paper. Finally, thanks to Paul Noordhof for many helpful conversations about delusions, and for titling the paper.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported from the Arts and Humanities Research Council [grant number AH/K003615/1); a European Research Council Consolidator [grant number 616358].

Notes on contributor

Ema Sullivan-Bissett is a Lecturer in Philosophy at the University of Birmingham, UK.

References

American Psychiatric Association (2013). *Diagnostic statistical manual of mental disorders* (Fourth edition, Text Revision (DSM-V-TR)).

Antonovsky, A. (1987). *Unravelling the Mystery of Health*. San Francisco: Jossey-Bass.

Bayne, T., & Pacherie, E. (2004a). Bottom-up or top-down: Campbell's rationality account of monothematic delusions'. *Philosophy, Psychiatry, & Psychology, 11*, 1–11.

Bayne, T., & Pacherie, E. (2004b). Experience, belief, and the interpretative fold. *Philosophy, Psychiatry, & Psychology., 11*, 81–86.

Bayne, T., & Pacherie, E. (2005). In defence of the doxastic conception of delusions. *Mind and Language, 20*, 163–188.

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

Bermúdez, José L. (2001). Normativity and rationality in delusional psychiatric disorders. *Mind & Language, 16*(5), 457–493.

Bortolotti, L. (2010). *Delusions and other irrational beliefs*. Oxford: Oxford University Press.

Bortolotti, L. (2012). In defence of modest doxasticism about delusions. *Neuroethics, 5*, 39–53.

Bortolotti, L. (2016). Epistemic benefits of elaborated and systematized delusions in Schizophrenia. *British Journal for the Philosophy of Science, 67*(3), 879–900.

Bortolotti, L. (2015). The epistemic innocence of motivated delusions. *Consciousness and Cognition, 33*, 490–499.

Bortolotti, L., & Broome, M. R. (2008). Delusional beliefs and reason giving. *Philosophical Psychology, 21*, 821–841.

Bortolotti, L., & Sullivan-Bissett, E. (forthcoming). The epistemic innocence of clinical memory distortions. *Mind & Language*. doi:10.1111/mila.12175

Broome, M. R. (2004). The rationality of psychosis and understanding the deluded. *Philosophy, Psychiatry, & Psychology, 11*, 35–41.

Butler, P. V. (2000). Reverse Othello syndrome subsequent to traumatic brain injury. *Psychiatry, 63*, 85–92.

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

Coltheart, M. (2005). Conscious experience and delusional belief. *Philosophy, Psychiatry & Psychology, 12*, 153–157.

Coltheart, M. (2010). The neuropsychology of delusions. *Annals of the New York Academy of Sciences, 1191*, 16–26.

Coltheart, M., Langdon, R., & McKay, R. (2007). Schizophrenia and monothematic delusions. *Schizophrenia Bulletin, 33*, 642–647.

Coltheart, M., Menzies, P., & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive Neuropsychiatry, 15*, 261–287.

Conrad, K. (1958). *Die beginnende Schizophrenie*. Stuttgart, Germany: Thieme Verlag.

Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R. F., Shanks, D. R., Robbins, T. W., … 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*, 345–369.

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.

Davies, M., Coltheart, M., Langdon, R., & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry, & Psychology, 8*, 133–158.

Ellis, H. (1998). Cognitive neuropsychiatry and delusional misidentification syndromes: An exemplary vindication of the new discipline. *Cognitive Neuropsychiatry, 3*, 241–258.
Fineberg, S. K., & Corlett, P. R. (2016). The doxastic shear pin: Delusions as errors of learning and memory. *Cognitive Neuropsychiatry, 21*(1), 73–89.
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*, 194–201.
Gerrans, P. (2001). Delusions as performance failures. *Cognitive Neuropsychiatry, 6*, 161–173.
Gerrans, P. (2002). A one-stage explanation of the cotard delusion. *Philosophy, Psychiatry, & Psychology, 9*, 47–53.
Gerrans, P. (2014). *The measure of madness*. Cambridge, MA: MIT Press.
Greenawalt, K. (1986). distinguishing justifications from excuses. *Law and Contemporary Problems, 49*, 89–108.
Hohwy, J., & Rosenberg, R. (2005). Unusual experiences, reality testing and delusions of alien control. *Mind and Language, 20*, 141–162.
Humphrey, N. (1995). *Soul searching*. London: Chatto and Windus.
Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgments in deluded and non-deluded subjects. *The Quarterly Journal of Experimental Psychology Section A, 40A*, 801–812.
Jaspers, K. (1963). *General psychopathology*. J. Hoenig, & M. W. Hamilton (Trans.). Manchester, NH: Manchester University Press.
Lansky, M. R. (1977). Schizophrenic delusional phenomena. *Comprehensive Psychiatry, 18*, 157–168.
Letheby, C. (2015). The epistemic innocence of psychedelic states. *Consciousness and Cognition, 39*, 28–37.
Maher, B. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology, 30*, 98–113.
Maher, B. (1988). Anomalous experience and delusional thinking: The logic of explanations. In T. Oltmanns & B. Maher (Eds.), *Delusional beliefs* (pp. 15–33). Hoboken, NJ: Wiley.
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*, 1–22.
Maher, B. (2006). The relationship between delusions and hallucinations. *Current Psychiatry Reports, 8*, 179–183.
McHugh, C. (2012). The truth norm of belief. *Pacific Philosophical Quarterly, 93*, 8–30.
McKay, R. (2012). Delusional inference. *Mind and Language, 27*, 330–355.
Mishara, A. L. (2010). Klaus Conrad (1905–1961): Delusional mood, psychosis, and beginning schizophrenia. *Schizophrenia Bulletin, 36*, 9–13.
Miyazono, K., Bortolotti, L., & Broome, M. (2014). Prediction-error and two-factor theories of delusion formation: Competitors or allies? In N. Galbraith (Ed.), *Aberrant beliefs and reasoning* (pp. 34–54). London: Psychology Press.
Noordhof, P., and Sullivan-Bissett, E. (manuscript). *The clinical significance of anomalous experience in the explanation of monothematic delusion*.
Puddifoot, K. (2017). Dissolving the epistemic/ethical dilemma over implicit bias. *Philosophical Explorations, 20*(sup1), 73–93.
Raskin, D. E., & Sullivan, K. E. (1974). Erotomania. *American Journal of Psychiatry, 131*, 1033–1035.

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.

Smith, B., Fowler, D., Freeman, D., Bebbington, P., Bashforth, H., & Garety, P. (2006). Emotion and psychosis: Links between depression, self-esteem, negative schematic beliefs and delusions and hallucinations. *Schizophrenia Research, 81*, 181–188.

Stone, T., & Young, A. W. (1997). Delusions and brain injury: The philosophy and psychology of belief. *Mind and Language, 12*, 327–364.

Sullivan-Bissett, E. (2015). Implicit bias, confabulation, and epistemic innocence. *Consciousness and Cognition, 33*, 548–560.

Sullivan-Bissett, E. (manuscript). *Unimpaired abduction to alien abduction: Lessons on delusion formation*.

Young, A., & Leafhead, W. (1996). Betwixt life and death: Case studies of the Cotard delusion. In P. Halligan & J. Marshall (Eds.), *Methods in adness: Case studies in cognitive neuropsychiatry* (pp. 147–171). Hove: Psychology Press.