Self-Deception, Delusion and the Boundaries of Folk Psychology *

Lisa Bortolotti [†]
l.bortolotti@bham.ac.uk

Matteo Mameli [††]
matteo.mameli@kcl.ac.uk

ABSTRACT

To what extent do self-deception and delusion overlap? In this paper we argue that both self-deception and delusions can be understood in folk-psychological terms. "Motivated" delusions, just like self-deception, can be described as beliefs driven by personal interests. If self-deception can be understood folk-psychologically because of its motivational component, so can motivated delusions. Non-motivated delusions also fit (to a large extent) the folk-psychological notion of belief, since they can be described as hypotheses one endorses when attempting to make sense of unusual and powerful experiences. We suggest that there is continuity between the epistemic irrationality manifested in self-deception and in delusion.

1. Introduction

1.1. Self-deception

In a fairly uncontroversial characterisation, self-deception involves beliefs that are acquired and maintained in the face of strong counter-evidence and that are motivated by desires or emotions (Deweese-Boyd, 2010). Self-deception is thought to be a widespread phenomenon in the general (non-clinical) population. Here is an example of self-deception. In spite of having at her disposal evidence to the contrary, Sylvia believes that she failed the driving test because the examiner was prejudiced against female drivers. Her belief responds to the need of preserving a positive image of herself as a competent driver. Here is another example. In spite of having at her disposal evidence that powerfully indicates that her son robbed a bank, Janet still believes that he is innocent. Her belief protects her from the acknowledgement of a truth (that her son is guilty) that is painful for her to accept.

There are two opposed philosophical accounts of self-deception. According to the traditional account, self-deception is due to the *doxastic conflict* between the false belief one acquires ("I failed the test because the examiner was prejudiced against female drivers") and the true belief one denies ("I failed the test because I drove badly").

According to the rival account, self-deception is due to *biased treatment of evidence*: there is a bias against considering or gathering evidence for the true belief. Sylvia never acquires the belief that failing the test was due to her poor driving, because she neglects evidence that points in that direction.

In the doxastic conflict account of self-deception, one has two contradictory beliefs, but is aware of only one of them, because one is motivated to remain unaware of the other (e.g., Davidson, 1982; 1986). On this view, when one deceives oneself, one believes a true proposition ("I failed the driving test because I drove badly") and acts in such a way as to cause oneself to believe the negation of that proposition ("I failed the test not because I drove badly but because the examiner was prejudiced against female drivers").

Doxastic conflict is problematic for two reasons. First, it involves accepting that one can believe a proposition and its negation at the same time, and some philosophers think that this is impossible (leading to the *static* paradox of self-

deception). Second, it suggests that one can intend to believe something that one knows to be false – and thus be the perpetrator and victim of a deceitful strategy all at once (leading to the *dynamic* paradox of self-deception). The solution some traditionalists offer for these puzzles consists in postulating *mental partitioning*. According to Davidson, one can have two mutually contradictory beliefs as long as one does not believe their conjunction. The idea is that each of the two beliefs is in a different compartment or partition of the mind, and this prevents the subject from recognising and eliminating the inconsistency.

If this account of self-deception prevails, the scope for identifying an area of overlap between self-deception and delusion is limited, as many delusions (those that are not "motivated") cannot be plausibly characterised as the simultaneous holding of two contradictory beliefs. That said, compartmentalisation can be observed in many people with delusions, when one's delusional belief is insulated from one's other beliefs that conflict with it.

A more revisionist solution to the puzzles generated by the doxastic conflict view leads to endorsing the competing account of self-deception. This account emphasises the differences between deceiving another and deceiving oneself. In the latter case, when the deceiver and the deceived are the same individual, deception need not be intentional, and the deceiver need not believe the negation of the proposition that she is causing the deceived to believe. If Sylvia wanted to deceive her father about the reason why she failed the driving test, the conditions for her deceiving him would be that she knows that she failed the test because she drove badly, but she intends to make her father believe otherwise. Self-deception works differently. Sylvia deceives herself if she genuinely comes to believe that she is not to blame for failing the test.

Al Mele argues that the conditions for self-deception are as follows. First, one's belief is false. Second, one treats the evidence relevant to the truth of the belief in a motivationally biased way. Third, this biased treatment of the evidence is what causes one to acquire the false belief. And finally, the evidence available to one at the time of acquiring the belief lends better support to the negation of one's belief than to the belief one acquires (Mele, 2001, pp. 50–51).

The ways in which the treatment of evidence can be motivationally biased are varied: one might misinterpret the available evidence, focus selectively on

those aspects of the available evidence that support one's belief, or actively search for evidence that supports one's belief, without also searching for evidence that disconfirms it (Mele, 2009). Motivationally biased treatment of evidence is not just relevant to the acquisition of the false belief, but also to its maintenance. One holds on to the false belief because one keeps neglecting some of the relevant evidence.

This deflationist approach is explanatory, and avoids the so-called paradoxes of self-deception. More importantly for our purposes here, the approach highlights the continuity between the phenomenon of self-deception and other instances of epistemic irrationality in ordinary beliefs and in delusions.

1.2. Delusion

According to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV, APA, 2000) and to the dominant theory of delusion formation in cognitive psychology (Coltheart, 2005), a delusion is a belief held with conviction and rarely challenged or revised. Delusional beliefs are typically implausible and unsupported by evidence (Bortolotti, 2010). As clinical delusions are symptoms of schizophrenia, dementia, and other psychiatric disorders, it is also important to add that they tend to disrupt day-to-day functioning (McKay, Langdon & Coltheart, 2005).

The content of some delusions (e.g., delusions of jealousy and of persecution) can be mundane and not dissimilar from that of false beliefs that we routinely find in the non-clinical population. Other delusions have more bizarre content. The Cotard delusion, for instance, is the belief that one is dead or disembodied. Delusions of infestation involve believing that insects are crawling under one's skin.

All delusions are currently thought to have an organic cause and are explained in neuropsychological terms, by reference to brain damage, perception failures, reasoning biases and cognitive deficits. But the formation of some delusions is also likely to include motivational factors. In a variety of anosognosia, people may fail to acknowledge the paralysis of a limb. This denial can be seen as a defence mechanism: one comes to believe that one's

arm is not paralysed because it is too hard to acknowledge that one permanently lost the use of one's arm.

There are similarities in the surface features of self-deception and motivated delusions – both phenomena typically involve beliefs that are badly supported by the evidence and that conflict with one's other beliefs or attitudes. Moreover, in both cases the beliefs are strikingly resistant to counterevidence. Further similarities can be found in the function of the beliefs: they serve to either preserve positive emotions, deny unpleasant or disturbing facts, or satisfy some other pressing psychological need.

Given what we know about self-deception and delusion, there are at least two features that distinguish the two: (a) whereas in self-deception beliefs are *always* motivated, not all delusional beliefs are motivated; (b) whereas delusions are symptoms of psychiatric disorders, are accompanied by other symptoms, and typically impair functioning, cases of self-deception are widespread in the non-clinical population.

Let us examine some of the differences and similarities in more detail.

2. The overlap between self-deception and delusion

There is no consensus on whether self-deception and delusion significantly overlap.[1] McKay and colleagues adopt the following approach to the issue:

> [S]ome (perhaps all?) delusional states may arise without self-deception, via processes that are not remotely motivated. [...] Conversely, self-deception may occur in a benign manner such that the resulting doxastic states do not sufficiently disrupt functioning to warrant the label *delusion*. (McKay et al., 2005, p. 315)

In this section, we will consider the notion of motivation as it applies to delusions and assess one interpretation of the view that self-deception is somehow more "benign" than delusion.

---

[1] See Mele 2009, Levy 2009, Davies 2009.

2.1. Motivation

Some delusions have been described as extreme cases of self-deception. They are considered cases of self-deception because they seem to have a defensive function. [2] They are considered extreme because they tend to disrupt functioning to a greater extent than standard cases of self-deception, and to result in the endorsement of more implausible and more tenacious beliefs.

One delusion that seems to fit this description is that of a delusion named "reverse Othello syndrome" which is the opposite of a delusion of jealousy. It consists in believing (incorrectly) that one's partner is faithful and in obstinately refusing to believe the contrary. The belief can plausibly be regarded as part of a defence mechanism against the suffering that the acknowledgement of the infidelity of one's partner would cause. [3]

Another example is anosognosia, the denial of illness. One well-known case is that of a woman (FD) who suffered from a right hemisphere stroke causing left hemiplegia (Ramachandran, 1996). FD could not move without a wheelchair and could not move her left arm. But when she was asked whether she could walk and engage in activities that require the use of both hands (such as clapping), she claimed that she could.

Vilayanur S. Ramachandran puts forward an explanation for this sort of cases. Behaviours giving rise to confabulations and delusions are an exaggeration of normal defence mechanisms which have an adaptive function. They allow one to preserve a positive self-image in the face of threatening negative events. The mind aims at maintaining a coherent system of beliefs that can guide behaviour. In normal subjects, the left hemisphere produces confabulatory explanations aimed at preserving the *status quo* ("My wife still loves me"; "My arm still moves"), but the right hemisphere detects discrepancies between the hypotheses generated by the left hemisphere and reality as it is perceived and it forces a revision of the belief system.

---

[2] See Ramachandran & Blakeslee 1998, Hirstein 2005.
[3] For a detailed description of one such case see Butler 2000 and the discussion by McKay et al. 2005, p. 313.

In patients with reverse Othello syndrome and anosognosia, the discrepancy detector in the right hemisphere malfunctions. A man suffering from the reverse Othello syndrome, for instance, claimed that his partner was faithful to him, whereas she had left him some time before (Butler, 2000). Thus, he failed to revise his belief in his partner's fidelity. In anosognosia, patients deny their own impairments even if they cannot help experiencing the effects of such impairments. In a conversation reported by Ramachandran (1996), FD asserted that her left arm was pointing at the doctor's nose, whereas her arm laid motionless.

Cases like these seem to support the claim that some delusions are motivated in much the same way as instances of self-deception are. It is prima facie plausible to regard delusions such as the reverse Othello syndrome and anosognosia as cases of self-deception, although the question can only be settled once we agree on an account of self-deception and find that it does fit the behavioural manifestations and the causal history of the beliefs. There are other delusions that deliver a boost to self-esteem: in erotomania, one believes that a person of higher status loves them (in secret); in delusions of grandeur, one believes to be a genius (unbeknownst to others); and delusions of persecutions often explain away instances of personal failure. For instance, a man can believe that he was fired because his colleagues conspired against him, whereas he was fired for his incompetence. Such delusions can also qualify as cases of self-deception in some circumstances. That said, it is important to stress that, even when motivational factors contribute to the formation of delusions, their presence is not sufficient to give rise to the delusion. Other factors (e.g., perception failures, brain damage, cognitive deficits, reasoning biases) need to be in place.

In addition, it is difficult to find any plausible role for motivational factors in the genesis of delusions such as the Cotard delusion, the belief that one is dead or disembodied. This delusion does not have an obvious adaptive function, and there is no fundamental role for motivational biases in the explanation of how the subject comes to hold or retain the delusional belief. Thus, the overlap

between self-deception and delusion can only be a partial one.[4] Motivational factors contribute to the formation of some but not all delusions, and only some delusions can be plausibly seen as the product of a psychological defensive mechanism. That said, there are still some interesting questions to answer. When delusions are motivated, are they *extreme* cases of self-deception? Are delusions in general more puzzling, less understandable, than standard cases of self-deception?

2.2. The boundaries of folk psychology

Recently, different conceptions of the relationship between delusion and self-deception have emerged. According to Keith Frankish (2011), both delusion and self-deception can be described by using the folk-psychological notion of belief, as long as the existence of different *types* of beliefs (roughly, behavioural dispositions and policies) is acknowledged. On his account, delusions and self-deception are continuous and motivational factors can contribute significantly to the formation of (at least some) delusions.

> Perhaps patients adopt delusions because they answer some emotional or other psychological need, rather than because they are probable. (Frankish, 2011)

Andy Egan (2009) also maintains that delusion and self-deception are alike, but takes the opposite line: *neither* can be accounted for satisfactorily by using the folk-psychological notion of belief. He argues that both delusion and self-deception should be regarded as *in-between* states. They represent how the agent takes things to be, and in this respect they are similar to beliefs. But they also convey how the agent wants things to be, and in this respect they are similar to desires. Egan suggests that they may be "besires", mental states that display at once features typical of beliefs and features typical of desires.[5]

---

[4] See McKay et al. 2005 and Davies 2009.
[5] Maura Tumulty (2011) and Eric Schwitzgebel (2011) also develop an account of delusions as in-between states.

In contrast to Frankish and Egan, Dominic Murphy (2011) highlights the discontinuity between delusion and self-deception. He maintains that instances of self-deception are understandable from a folk-psychological perspective, whereas delusions are not. We want to concentrate on Murphy's view here.

Murphy uses the following example to argue that self-deception is understandable from a folk-psychological perspective and to argue for the existence of a discontinuity between self-deception and delusions.

> It is easy to imagine parents who refuse to acknowledge that their child is guilty of a heinous crime, despite sufficiently overwhelming evidence to convince everyone else that the guilty verdict is the right one. Let's suppose that the child is guilty, and that everyone else believes this because it is the correct inference to make given the evidence. The mother of the guilty man has no relevant evidence not possessed by others, but the cost to her of admitting her child's guilt is too great. (Murphy, 2011)

In line with the accounts of self-deception we cited earlier, Murphy claims that self-deception involves having beliefs that carry emotional commitment and are fixed by personal interests rather than by a careful consideration of the available evidence. According to Murphy, such personal interests offer an acceptable explanation of both the conflict between belief and evidence and the "rigidity" of the belief. Murphy recognises that the epistemologist would consider desire-driven beliefs as not rational, but he thinks that they are an understandable manifestation of human nature.

Typically, delusions are also poorly supported by evidence and scarcely responsive to counter-evidence, but these features cannot (always or entirely) be explained by the influence of desires. Moreover, the content of delusions is somehow more "unbelievable" than the contents we routinely deceive ourselves about. There is nothing absurd in believing that a man is innocent, even if the belief is clearly false given the evidence at one's disposal, but there is something deeply unsettling about the content of many delusions. We take this to be the point of Murphy's next example.

> Let's consider another case, this time the (real) case of a person I'll call Ed. Ed was sleeping rough, and heard a tree in a park tell him that the park was a good place to stay. So Ed settled down for the night in the park. But a little later, the sprinklers in the park erupted and Ed was drenched. Thereupon Ed heard the

> tree tell him that it was very sorry: trees like to be watered, and the tree had not
> understood that Ed would not appreciate a good soaking. Ed accepted the
> tree's apology and went on his way. [...] Ed's traffic with trees is evidence of
> something mentally abnormal about him. (Murphy, 2011)

Murphy argues that delusion (but not self-deception) remains mysterious from
a folk-psychological perspective.

> Ed [...] seems incomprehensible in folk terms; he is a suitable case for
> treatment. Delusions, I suggest, are attributed [...] when we run out of the
> explanatory resources provided to us by our folk understandings of how the
> mind works. (Murphy, 2011)

If motivational factors can contribute to the formation of at least some
delusions, then Murphy's view about the discontinuity between self-deception
and delusion *in general* is problematic. If the fact that a desire motivates a
belief is sufficient for the folk-psychological understandability of such belief,
no matter how impervious to evidence the belief turns out to be or how
implausible, then only those delusions that are not motivated defy folk-
psychological explanation. This view is compatible with the claim that, among
delusions, those that are not motivated lack the folk-psychological
understandability that both instances of self-deception and motivated delusions
have. On this account, delusions occurring in anosognosia and the reverse
Othello syndrome are amenable to folk-psychological explanation, while the
Cotard delusion and delusions of infestation, as well as Ed's delusion about the
talking tree, are not.

One may want to deny that the fact that a desire motivates a belief is
sufficient for the folk-psychological understandability of such belief. On this
view, even motivated delusions are discontinuous with self-deception because
the role of motivational factors in their formation cannot provide an adequate
explanation of the bizarre content of the resulting beliefs or of their
imperviousness to counterevidence. Thus, a mother's love for her son can
explain why she refuses to believe that he committed a crime, but one's desire
not to be paralysed cannot explain the denial of the paralysis.

We find this latter view implausible. The denial of a serious physical
impairment can surely be explained, at least in part, by reference to the relevant
motivational states and is therefore folk-psychologically understandable, just

like the refusal of a mother to acknowledge that her son is guilty of a heinous crime. Both beliefs seem to have a defensive function and respond to a psychological need. There are many relevant similarities between the two cases epistemically, such as neglect and misinterpretation of evidence, implausibility and tenacity of the belief. Even considering the role of cultural norms, there seems to be no important difference: just like the acknowledgement that one's son is guilty of a crime, the acknowledgement of a serious and permanent impairment is something people have a reason to avoid. In both cases, from a folk-psychological perspective, it is not surprising that people sometimes believe what they would like to be true.

Let us now consider the more modest claim that *non-motivated* delusions are not understandable within the framework of folk psychology. We think this claim should be resisted too. In his analysis, Murphy focuses on the agent's reasons for her treatment of evidence. One could say that in self-deception (and in motivated delusions) evidence is neglected or misinterpreted *for a reason* (e.g., personal interests that are culturally recognisable) but in non-motivated delusions evidence is neglected or misinterpreted for no reason. When one considers the question whether one's right leg is paralysed, one might neglect to consider as relevant evidence the fact that one can no longer climb stairs. This evidence is neglected or discounted due to one's desire to believe that one's right leg is not paralysed. When one considers the question whether one is disembodied, one might neglect to consider whether one can move, talk and feel. This evidence is neglected or discounted but it is not clear why, as there seems to be no interest in believing that one is disembodied.

An issue that needs addressing is how demanding we take the folk-psychological notion of belief to be. Murphy claims that in the case of the mother deceiving herself about the innocence of her son the belief is in some respects faulty on epistemic grounds (i.e., not supported by or responsive to evidence) but not necessarily irrational. The belief does not conflict with behavioural generalisations that belong to our folk theory of the mind in an extended sense.

> These resources [provided to us by our folk understanding of the mind] do not just include folk psychology in the narrow sense of theory of mind, but a much richer body of beliefs and expectations about the role of hot cognition and personal interests in fixing belief [...] and the role of culture in shaping

> people's assumptions about what counts as legitimate evidence. (Murphy, 2011)

If the folk-psychological notion of belief were very demanding, and required that all legitimate beliefs be supported by and responsive to evidence, then both self-deception and delusion would fail to count as instances of belief. After all, according to the demanding interpretation of the folk-psychological notion of belief, desires do not interfere directly in the formation of beliefs at the expense of evidence – there are no *besires* in old-school folk psychology. Mental states are beliefs in virtue of their relationship to other beliefs (e.g., inferential relations), their relationship to behaviour (e.g., action-guiding potential), and especially their relationship to evidence. A mental state that is formed on the basis of partial evidence and that is scarcely responsive to new evidence would fall short of being a belief in a rigid, uncompromising framework.

But the folk-psychological notion of belief seems to be compatible both with the idea that in some cases desires play a role – even a direct role – in the formation of beliefs and with the idea that there are irrational beliefs. Folk-psychology can allow for the case of someone who believes that she has become disembodied after her experience of herself in relation to the rest of the world suddenly changed. After all, the relationship between unusual experiences and bizarre delusions is the relationship of evidence supporting a belief. Folk-psychology can also allow for Ed's delusional belief that the tree talked to him. The delusion is not without a reason if (we are elaborating the original example here) Ed heard voices in the park but saw nobody around. There are probably no *good* reasons to suppose that a tree is talking, but we do not need *good* reasons to establish the comparative claim with self-deception. Wanting one's son to be innocent is not a good reason to believe that he is.

In some respects, non-motivated delusions seem to be even more typical cases of belief than the case of the mother refusing to accept that her son is guilty. Not only instances of Cotard delusion, delusions of infestations and Ed's belief in talking trees are likely to interact with other beliefs and to guide action, as standard beliefs do, but such mental states are there to make sense of weird experiences with specific contents, experiences which would otherwise be inexplicable to those who are not acquainted with the form that psychotic symptoms can take. This is not the whole story, of course. Delusions are

irrational beliefs because they are not revised when counterevidence becomes available. But being scarcely responsive to some of the available evidence is one of the features delusions have in common with cases of self-deception, so the continuity between the two phenomena is not compromised.

To sum up, if folk psychology can allow for beliefs formed in order to satisfy a desire, and for beliefs that are poorly supported by and scarcely responsive to evidence, then it can also account for delusions.

3. Epistemic irrationality

Philosophers explain the status of self-deception and delusion differently. As we saw, some suggest that they are types of beliefs and some suggest that they are in-between states, which share some features with beliefs and other features with imaginings or desires. We would like to suggest that both self-deception and delusion are beliefs that violate norms of epistemic rationality. This claim is consistent with accepted definitions of both delusion and self-deception, but in order to make the claim meaningful one needs to formulate a notion of epistemic rationality and to distinguish it from other notions of rationality.

There are (at least) three forms of rationality that apply to belief-like states: *procedural*, *epistemic* and *agential* rationality (Bortolotti, 2009). Procedural rationality concerns the relationship between a belief and one's other beliefs. A clear violation of procedural rationality is inconsistency among one's beliefs. Epistemic rationality concerns the relationship between a belief and the available evidence. A clear violation of epistemic rationality is hanging on to a belief that has been repeatedly challenged by reliable evidence. Agential rationality concerns the relationship between a belief and behaviour. A clear violation of agential rationality is acting in a way that conflicts with one's belief.

Delusion and self-deception may violate more than one set of norms, but they are typically beliefs at odds with the evidence. Norms of epistemic rationality govern the acquisition, maintenance and revision of beliefs. Epistemically irrational beliefs can be badly supported by one's initial evidence or scarcely responsive to evidence that becomes available at a later stage. Evidence in support of the hypothesis that if the sky is red at night, then the weather will be good on the following day ("Red sky at night; shepherds

delight") should be weighed up by a rational subject before she takes the hypothesis to be true. Further, if evidence against the hypothesis becomes available after the hypothesis has been endorsed, and this evidence is sufficiently powerful, robust and so on, then the rational subject should come to doubt the previously formed belief, suspend judgement until new evidence becomes available, or reject the belief altogether.

As we previously discussed, forming a hypothesis ("My son is not guilty", "My left arm can move", "Insects are crawling under my skin") that is not supported by all the available evidence is not necessarily problematic. What seems problematic is to endorse such hypothesis as a belief and to hang onto the belief in the face of evidence that openly conflicts with it. Suppose the son confesses the crime to his mother and she discounts his confession. Suppose the patient continues to believe that he is not paralysed after the doctor explains to him in no vague terms what his situation is. In these circumstances, if the hypothesis is not shaken by such challenges but crystallises into a tenacious belief, then something is amiss.

As you may remember, Murphy agrees that the mother's belief in the son's innocence is epistemically irrational, as it is not supported by the evidence. Murphy also thinks that the mother's behaviour is folk-psychologically understandable and that we would not consider it as irrational *tout court*. One way of making the point is that the mother's belief is epistemically irrational but it is pragmatically rational for her to have that belief, in the sense that her life would be worse (all things considered) if she gave up the false belief and acknowledged that her son is indeed guilty. What is interesting is that *some* delusions also seem to work in the same way. By definition (at least the DSM-IV definition), delusions are epistemically irrational beliefs, but it is *not always* pragmatically irrational to be delusional. Imagine you are in Ed's shoes. The alternative to believing that the tree just talked to you is to concede that you hear voices and something is seriously wrong with you.

Aikaterini Fotopoulou explains that after brain damage or memory loss, personal narratives can be disrupted, undermining people's sense of coherence. This is often associated with increased anxiety and depression. Despite their poor correspondence with reality, delusional and confabulatory beliefs represent attempts to define one's self in time and in relation to the

world. Thus, they are subject to motivational influences and they contribute to preserving one's identity (Fotopoulou, 2008, p. 542).

People with delusions and confabulations construct distorted or false self-conceptions. They may claim that they live in a different place from the one where they live, or that they have a different profession or a different family from the one they do. The personal narrative they construct is not «anchored and constrained by reality» (Fotopoulou, 2008, p. 548). These distortions are exaggerated by brain damage or memory loss and exhibit self-serving biases – people reconstruct and interpret events in a way that is consistent with their *desired* self-image.

For the sake of creating a coherent self-image, people enhance their life-stories. In dementia, amnesia, anosognosia, people revisit their present and their past and attempt to establish continuity between the conception they had of themselves before the accident, the memory loss, the illness, and the conception of themselves afterwards. In this reconstruction, people tend to preserve a positive image whenever possible. Maintaining coherence with the previous self-image and promoting a more positive self-image take priority over preserving accuracy. The preference for internal coherence over correspondence has consequences.

The obvious disadvantage is that losing touch with reality can create a gulf between the person with the delusion and the surrrounding social environment. In the most serious amnesic conditions there is often a lack of "shared reality" between confabulators and the people who were once closest to them, which can be very distressing for patients and their families (Fotopoulou, 2008, p. 560). In general, given that delusions are ill-grounded and often bizarrely false, people with delusions are not likely to be believed and taken seriously by others.

These observations on distorted memory and enhanced self-narratives in the clinical population affected by delusions and confabulations apply also to self-deception. In this respect, the oft-perceived gap between delusions as a clinical, pathological phenomenon and self-deception as a homely form of epistemic irrationality seems to shrink. Non-clinical subjects also tend to present their current selves in a way that is both coherent with their past, and largely favourable (Wilson & Ross, 2003), giving rise to common instances of self-deception. Self-deception can also result in a gulf between one's version of

reality ("The examiner was biased against female drivers", "My son is innocent") and the version of reality other people share and accept. The clinical case helps us realise that the development of self-narratives is always a *reconstructive* exercise, even when memory and reasoning are not seriously compromised.

> A self-conception is not just the set of facts we might learn about ourselves; it is an interpretation of these facts within which values are prioritized, emotions are labeled, and attitudes are endorsed or rejected. Importantly, the process of organizing what we know about ourselves into a self-conception is partly a creative or constructive process. (Tiberius, 2008, p. 116)

The fact that delusion and self-deception involve irrational beliefs does not mean that they bring no benefits at all. As previously suggested, delusion and self-deception may have some pragmatic benefits. They protect the subject from undesirable truths, keep anxiety and depression at bay, and help maintain a coherent sense of self (Bortolotti & Cox, 2009). They allow people to keep constructing self-narratives when personal information is not available, and to construct self-narratives that are more positive than the evidence suggests, preserving self-esteem in the face of serious set-backs.

4. Conclusion

In this paper, we revisited a topic that has engaged philosophers of mind in recent years, the potential overlap between self-deception and delusion. Our purpose was to show that, although the two phenomena are distinct, there is considerable continuity between them. We argued against the claim that delusion does not fit the folk-psychological notion of belief, whereas self-deception does. If instances of self-deception can be understood folk-psychologically, then delusions can too.

By appealing to the notion of epistemic irrationality, we suggested that in self-deception and delusion the relationship between belief and evidence is unhealthy, which causes delusional and self-deceiving people to form inaccurate accounts of themselves and of the events that concern them. As a result, the delusional and the self-deceived may reject the view of themselves or of reality that people around them share in order to preserve a positive and coherent sense of self.

REFERENCES

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision.*

Bayne, T., & Fernandez, J. (Eds.) (2009). *Delusions and Self-Deception: Affective Influences on Belief Formation.* Hove: Psychology Press.

Bortolotti, L. (2009). *Delusions and Other Irrational Beliefs.* Oxford: Oxford University Press.

Bortolotti, L. (2010). Delusion. In E.N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Fall 2010 Edition)* <http://plato.stanford.edu/archives/fall2010/entries/delusion/>.

Bortolotti, L., & Cox, R. (2009). Faultless ignorance: strengths and limitations of epistemic definitions of confabulation. *Consciousness & Cognition, 18*(4), 952– 965.

Butler, P.V. (2000). Reverse Othello syndrome subsequent to traumatic brain injury. *Psychiatry: Interpersonal and Biological Processes, 63*(1), 85– 92.

Coltheart, M. (2005). Delusional belief. *Australian Journal of Psychology, 57*(2), 72–76.

Davidson. D. (1982). Paradoxes of irrationality. In R. Wollheim & J. Hopkins (Eds.), *Philosophical essays on Freud.* Cambridge: Cambridge University Press, 289–305.

Davidson, D. (1986). Deception and Division. In J. Elster (Ed.), *The Multiple Self.* Cambridge: Cambridge University Press, 79–92.

Davies, M. (2008). Delusion and Motivationally Biased Belief: Self-Deception in the Two-Factor Framework. In T. Bayne, & J. Fernandez (Eds.) *Delusions and Self-Deception: Affective Influences on Belief Formation.* Hove: Psychology Press, 71–86.

Deweese-Boyd, I. (2010). Self-deception. In E.N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy (Fall 2010 Edition),*

<http://plato.stanford.edu/archives/fall2010/entries/self-deception/>.

Egan, A. (2008). Imagination, Delusion, and Self-Deception. In T. Bayne, & J. Fernandez (Eds.), *Delusions and Self-Deception: Affective Influences on Belief Formation*. Hove: Psychology Press, 263–280.

Frankish, K. (2011). Delusions, Levels of Belief, and Non-doxastic Acceptances. *Neuroethics*, DOI: 10.1007/s12152-011-9123-7.

Hirstein, W. (2005). *Brain fiction: self-deception and the riddle of confabulation*. Cambridge, MA: MIT Press.

Levy, N. (2009). Self-deception without thought-experiments. In T. Bayne, & J. Fernandez (Eds.), *Delusions and Self-Deception: Affective Influences on Belief Formation*. Hove: Psychology Press, 227–242.

McKay, R., Langdon, R., & Coltheart, M. (2005). "Sleights of mind": Delusions, defences and self-deception. *Cognitive Neuropsychiatry*, *10*(4),305–326.

Mele, A. (2001). *Self-Deception Unmasked*. Princeton: Princeton University Press.

Mele, A. (2009). Self-deception and delusion. In T. Bayne and J. Fernandez (Eds.), *Delusions and Self-Deception: Affective Influences on Belief Formation*. Hove: Psychology Press, 55–70.

Murphy, D. (2011). The folk epistemology of delusions. *Neuroethics*, DOI: 10.1007/s12152-011-9125-5.

Ramachandran, V.S. (1996). The evolutionary biology of self-deception, laughter, dreaming and depression: some clues from anosognosia. *Medical Hypotheses*, *47*(5), 347–362.

Ramachandran, V.S., & Blakeslee, S. (1998). *Phantoms in the Brain: human nature and the architecture of the mind*. London: Fourth Estate.

Schwitzgebel. E. (2011). Mad belief? *Neuroethics*, DOI: 10.1007/s12152-011-9127-3.

Tiberius, V. (2008). *The Reflective Life: Living Wisely With Our Limits*. New York: Oxford University Press.

Tumulty, M. (2011). Delusions and not-quite beliefs. *Neuroethics*, DOI: 10.1007/s12152-011-9126-4.

Wilson, A., & Ross, M. (2003). The identity function of autobiographical memory: Time is on our side. *Memory*, *11*(2), 137–149.