Chapter 15

Explaining pathologies of belief

Anne M. Aimola Davies and Martin Davies

Abstract

The two-factor framework for explaining delusions is developed in a way that promises reasonable coverage without overgeneralization. We propose that heterogeneity in explanations of delusions can be conceived as parametric variation within the two-factor framework and we suggest several parameters. In three ways, we confront the fact that the second factor in the two-factor framework, a presumed impairment of belief evaluation, has been poorly specified in terms of cognitive function. First, an *a priori* task analysis suggests that belief evaluation involves working memory and executive processes of inhibition. Second, we review experimental and neuroimaging studies of the belief-bias effect in the context of dual-process accounts of reasoning. The results can be interpreted as supporting the proposal that the second factor in the explanation of delusions is an impairment of working memory or executive function with a neural basis in damage to the right frontal region of the brain. Finally, we present results from a study of cognitive impairments following stroke to support our proposal in the case of anosognosia considered as a delusion.

15.1 Introduction

In a case of delusion, belief goes wrong. A delusion is a belief that not only departs from the norms of truth and knowledge, but also is unresponsive to considerations of plausibility and evidence. A delusion is: 'A false belief . . . that is firmly sustained despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary' (*DSM-IV-TR*, 2000, p. 821). Delusions are pathologies of belief.

This notion of a pathology of belief can usefully be distinguished from a conception of pathological belief or doubt, that figures in some recent work in epistemology (Pryor, 2004). Having evidence to doubt the proposition that there is an external world, for example, could undermine a subject's justification, based on perceptual

experience, to believe the proposition that there is a table in front of him. In this context, it is important to distinguish between a doubt that is really supported by evidence and a doubt that the subject wrongly takes to be supported by evidence. It is also important – especially for the purposes of this chapter – to distinguish between a doubt that the subject takes to be supported by evidence (whether rightly or wrongly) and a pathological doubt – that is, a doubt that the subject knows to be unjustified but cannot help having. In a case of pathological doubt, the subject is beset by doubt but can offer no grounds for the doubt. Similarly, we can say that, in a case of pathological belief, the subject is *beset by belief* but can offer no grounds for the belief.

It is plausible that some cases of delusion are examples of pathological belief in this sense. Jaspers (1963) conceived of primary delusional beliefs in this way and these cases may be theoretically important. But there are surely other cases of delusion in which the subject does offer grounds for his or her belief, reasons that, at least from the subject's point of view, speak in favour of the belief. Conversely, there are imaginable cases of pathological belief that are not cases of delusion. A subject might, in principle, be beset by a belief that happens to be plausible, true, shared by other people, and consistent with available evidence. In short, when we say that delusions are pathologies of belief – that is, cases where belief goes wrong – we do not mean that delusions are pathological beliefs – that is, beliefs for which the subject can offer no grounds.[1]

15.2 **Anosognosia as a pathology of belief**

Anosognosia is a failure to acknowledge illness or impairment. Patients with anosognosia for their motor impairments following right-hemisphere stroke fail to acknowledge, and may outright deny, that they can no longer raise their left arm or move their left leg. Patients with anosognosia for the consequences of their motor impairments fail to appreciate their limited ability to carry out activities of daily living. They may insist, quite unrealistically, that they could live at home and care for themselves unaided.

Berti and colleagues (Berti *et al*., 1998) describe the case of an 80-year-old woman, CC, who suffered left-side paralysis following a stroke that caused damage to fronto-parietal subcortical regions of the right hemisphere (1998, p. 27). When examined during the 2 months after her stroke, patient CC did not acknowledge her motor impairments, even when they were demonstrated to her. She not only insisted that she could move her left arm but also maintained that she was moving it in the period immediately after being asked to do so. She did, however, show some appreciation of the consequences of her impairments. When asked to rate how well she would perform

[1] Bortolotti and Broome (2008, in press) make use of the notion of *authorship* of a belief (Moran, 2001), conceived as 'the capacity to endorse the content of a belief and justify it with reasons' (2008, p. 822). They consider the question whether delusions are beliefs of which the subject is not the author and provide convincing examples of delusions that are not authored – cases in which the subject can offer 'no explanation or reason to believe that what they say is true' *(ibid*., p. 829). But they also describe cases in which subjects with delusions are 'able to defend the content of the beliefs they report' (p. 829).

if she had to carry out an everyday task (such as lifting a glass) using her right hand or her left hand, she gave high scores for the right hand but low scores for the left hand.

House and Hodges (1988) describe the case of an 89-year-old woman who suffered left-side paralysis following a stroke that damaged the right basal ganglia but spared cortical regions. She was confined to a wheelchair and dependent on assistance for activities of daily living such as washing, grooming, and dressing. When examined 6 months after her stroke, she acknowledged some weakness, particularly when her impairments were demonstrated to her. But she insisted that 'she would be able to walk, feed, and dress herself unaided, and even drive a car although "the left side might be a bit awkward"' (1988, p. 114).

In a study by the first author (Aimola, 1999; Maguire and Ogden, 2002), patient M3 was a 59-year-old man who suffered severe left-side motor impairments following a right-hemisphere stroke that damaged the parietal, frontal, and temporal lobes and basal ganglia. Nine months after his stroke, he was confined to a wheelchair and would sometimes acknowledge his impairments. But he did not appreciate the consequences of his impairments. While at home alone, he repeatedly tried to get out of his wheelchair and injured himself. He had to be placed in a nursing home for his own safety. Patient M6, a 57-year-old man, also had severe motor impairments in the acute stage following a right-hemisphere stroke that caused extensive damage to the parietal, frontal, and temporal lobes and basal ganglia. Three months after his stroke he had made a relatively good recovery and was able to walk, although the weakness of his left leg was still evident as he needed to use a cane. Eight functional tests of hemiplegia (Gialanella and Mattioli, 1992) revealed that activities of daily living were possible for patient M6 only with difficulty. Nevertheless, he insisted that he could leave the rehabilitation hospital, live at home, and generally care for the family, even though this proved clearly beyond him when he made short visits home.

These patients failed to acknowledge their motor impairments or failed to appreciate the consequences of those impairments (or both). They overestimated their abilities to move their left-side limbs or their abilities to carry out activities of daily living and they maintained their false beliefs in the face of abundant evidence about their real situations. Anosognosia is a delusion, a pathology of belief. In the mid-twentieth century, it was common to explain anosognosia as a case of motivated denial. Weinstein and Kahn (Weinstein and Kahn, 1950, 1951, 1953, 1955; Weinstein *et al.*, 1954) put forward an influential account of anosognosia as an expression of the drive to be well that is present in everyone. Since the drive is not expressed as anosognosia in everyone who suffers from motor impairments, they proposed that 'the occurrence of anosognosia is related to the pattern of the premorbid personality' (1950, p. 780).

In recent years, explanations of anosognosia as motivated denial have fallen from favour and explanations in terms of sensory, attentional, and cognitive deficits have been preferred. One factor in this change has been Bisiach and Geminiani's (1991) influential argument opposing interpretations of anosognosia as 'a defensive adaptation against the stress caused by the illness' (1991, p. 24).[2] More generally, the change

[2] For extended discussion of the possible role of motivation in anosognosia and other delusions, see Aimola Davies *et al.*, 2009; Davies, 2009; Mele, 2009.

of approach is consistent with the development of cognitive neuropsychiatry – the use of the methods of cognitive neuropsychology for understanding disorders that were previously regarded as psychiatric phenomena.

15.3 Cognitive neuropsychology and cognitive neuropsychiatry

Research in cognitive neuropsychology has two complementary aims. One is to use data from people with acquired disorders of cognition to constrain, develop, and test theories of normal cognitive structures and processes. The other is to use theories about normal cognition to help understand disorders of cognition that result from stroke or head injury (Coltheart, 1985; Humphreys, 1991). It follows from the aims of cognitive neuropsychology that 'the underlying construct is . . . a model of normal performance in some cognitive domain or other' (Halligan and Marshall, 1996, pp. 5–6).

Language is arguably the cognitive domain in which cognitive neuropsychological research has been most highly developed. Beginning from two papers by Marshall and Newcombe (1966, 1973), the cognitive neuropsychology of reading has yielded theoretical accounts of acquired disorders of reading (dyslexias) in terms of a model of the cognitive structures and processes implicated in normal reading of words aloud (Coltheart, 2006a). The model is highly articulated and, at least partly, computationally implemented (Coltheart, 2006b). Some of the processes for reading aloud draw on orthographic and phonological information stored in the lexicon and can therefore only be applied to real words. Other processes make use of letter–sound (more accurately, grapheme–phoneme) correspondence rules and can be applied to pronounceable letter strings whether or not they are real words. Selective damage to some components of the model can thus explain impaired reading of irregular (exception) words (e.g. 'pint' pronounced to rhyme with 'mint') while reading of regular words (e.g. 'print') and non-words (e.g. 'slint') is spared. Selective damage to other components can explain impaired reading of non-words while reading of real words, both regular and irregular, is spared.

15.3.1 Cognitive neuropsychiatry

Hadyn Ellis is credited with the first public use of the term 'cognitive neuropsychiatry' for the application of the methods of cognitive neuropsychology to psychiatric disorders (in October 1991; see David, 1993, p. 4; Coltheart, 2007, p. 1042). In 1996, Halligan and Marshall's edited volume, *Method in Madness: Case Studies in Cognitive Neuropsychiatry*, was published and the journal, *Cognitive Neuropsychiatry*, was launched. The journal editors note some changes of approach attendant on the shift from the more familiar territory of cognitive neuropsychology (David and Halligan, 1996, p. 2): 'We need to think of excesses as well as deficits; transient rather than stable phenomena; distortions and biases rather than striking quantitative or apparent qualitative differences.' Young (2000) also reviews a 'catalogue of problems to be faced' and suggests that 'advances in cognitive neuropsychiatry will be hard won . . . [but] well worth the effort' (2000, p. 69).

An important early work in cognitive neuropsychiatry is Ellis and Young's paper, 'Accounting for delusional misidentification' (1990). One reason that this work was 'promising from the start' was that 'they [Ellis and Young] had a fairly simple yet well-substantiated model of face recognition based on studies of normal and clinical subjects including cases of prosopagnosia [Bruce and Young, 1986]' (David, 1993, p. 4; see also Ellis, 1998). A model of normal face recognition is clearly important for understanding delusions of misidentification, such as the Capgras delusion. But the 'underlying construct' that is required whenever the methods of cognitive neuropsychology are applied to pathologies of belief is a model of the normal formation, evaluation, and revision of beliefs. Thus, one of the problems faced by cognitive neuropsychiatry – in comparison with the cognitive neuropsychology of reading, for example – is that we do not have an articulated, still less a computationally implemented, model of normal believing. Indeed, there may be reasons of principle why it is extremely difficult to understand belief formation in terms of the computational theory of mind (Fodor, 1983, 2000).

The lack of a model of normal belief presents a challenge for cognitive neuropsychiatry. Halligan and Marshall display an optimistic and constructive spirit, saying (1996, p. 8): 'One would none the less hope that theories of normal belief-formation will eventually cast light on both the content of delusions and on the processes whereby the beliefs came to be held.' They also say that 'it is unlikely that a unified theory of delusions will be forthcoming' (*ibid.*). This latter idea is developed by Stone and Young in a seminal contribution to inter-disciplinary theorizing about delusions, 'Delusions and brain injury: The philosophy and psychology of belief' (1997). They draw on the analogy between the cognitive neuropsychology of reading and the cognitive neuropsychiatry of delusions. Just as 'there can be different non-word reading deficits, resulting from the precise way in which spelling-to-sound conversion has been impaired', so also, 'there are different kinds of delusions. The precise nature of a delusion will depend, *inter alia*, upon the exact way in which the system supporting belief formation has been impaired' (1997, p. 331).

We agree that there are different kinds of delusions and that their explanations will be correspondingly different. Nevertheless, we propose that the explanations of a wide range of delusions exhibit a kind of unity. The explanations can be conceived in terms of parametric variation within a single explanatory framework, rather as natural languages can be conceived in terms of parametric variation within a single universal grammar (Chomsky, 1986).

15.3.2 The two-factor framework

Coltheart (2007, p. 1044) has proposed that, in order to explain any delusion, we need to answer two questions. First, where did the delusion come from? Second, why does the patient not reject the belief? This is the leading idea of the two-factor framework for explaining pathologies of belief. The first factor figures in the explanation of how the patient came to regard the false proposition as a salient and serious hypothesis and initially adopted the hypothesis as a belief. The second factor figures in the explanation of the patient's maintenance of the belief despite its implausibility and despite the evidence against it.

In line with the developing research programme of cognitive neuropsychiatry, the two-factor framework was initially presented as a schematic explanation for delusions of neuropsychological origin. In such cases, it is reasonable to expect that a first neuropsychological deficit will provide (at least part of) an answer to the question where the delusion came from and that a second deficit will explain why the patient does not subsequently reject the false belief. The explicitly neuropsychological development of the two-factor framework is thus a 'two-deficit account of delusional belief' (Coltheart, 2007, p. 1044).

The scope of the two-factor framework might gradually be extended from neuropsychological cases of monothematic delusion to include cases of delusion without apparent brain injury and, ultimately, the floridly elaborated delusional systems of some individuals with schizophrenia. But broader explanatory coverage requires less specific commitments concerning the nature of the explanatory factors. So there is a risk that the account will overgeneralize – perhaps, in the worst case, encompassing all false beliefs. Our aim is to develop the two-factor framework in a way that offers the prospect of reasonable coverage without overgeneralization.

15.4 The first factor: where did the delusion come from?

In most cases of delusion, the subject's false belief is new and also bizarre or exotic. The subject may say: 'This [the subject's left arm] is not my arm' (somatoparaphrenia; Halligan *et al*., 1995; Bottini *et al*., 2002) or: 'This [the subject's wife] is not my wife. My wife has been replaced by an impostor' (Capgras delusion; Capgras and Reboul-Lachaux, 1923; Edelstyn and Oyebode, 1999). The answer to the question where the delusion came from may appeal to the subject's explanation or interpretation of an anomalous experience (Maher, 1974, 1988, 1992). In the neuropsychological version of the two-factor framework, we assume that the anomalous experience arises from a first deficit. Coltheart describes it in this way:

> The patient has a neuropsychological deficit of a kind that could plausibly be related to the content of the patient's particular delusion – that is, a deficit that could plausibly be viewed as having prompted the initial thought that turned into a delusional belief.
>
> (2007, p. 1047)

It is assumed that the first deficit varies from delusion to delusion and may also vary from patient to patient with the same delusion.

Neither a neuropsychological deficit nor an anomalous experience can provide a complete answer to the question where the delusion came from. A delusion is a belief, but having a deficit or experience is not yet having a belief; it is not even having a hypothesis that could be adopted as a belief. A complete answer to the question will have to appeal to a processing stage that leads from deficit or experience to belief. This is the idea that the two-factor framework is also a three-stage framework (Aimola Davies *et al*., 2009).

15.4.1 Endorsement or explanation

If an anomalous experience figures in the answer to the question where the delusion came from, then the representational content of the experience may be close to the content of the delusion itself or it may be very different. Suppose, at one end of the

spectrum of possibilities, an experience *fully encodes* the content of the delusion. In this case, what is needed, to lead from anomalous experience to delusional belief, is just that the subject should take the experience at face value or *endorse* it (Bayne and Pacherie, 2004). That is, the subject should treat the experience as veridical. Plausibly, this is a default or prepotent doxastic response to perceptual experiences (Davies *et al.*, 2001, p. 153).[3]

Now suppose, at the other end of the spectrum of possibilities, that the representational content of an anomalous experience is *much less specific* than the content of the delusion to which it leads. For example, the experience might be a feeling of significance or a conscious sense that something has changed (Maher, 1999). There is a substantial gap between the inchoate sense that a limb lying beside my torso is different or not quite right and the belief that it is not my arm but someone else's, or between the sense that a person who looks like my wife is different or not quite right and the belief that she is not my wife but an impostor. In such cases, the processing stage that leads from experience to belief must involve substantive *explanatory* processes of hypothesis generation and confirmation.

Continuing with the explanationist option, suppose E is the evidence provided by an anomalous experience and that an explanatory hypothesis, H, is generated. If the probability of the evidence E given the hypothesis H, *Pr*(E/H), is greater than the prior probability of E, *Pr*(E), then Bayes's theorem, in the form:

$$\frac{Pr(\mathrm{H/E})}{Pr(\mathrm{H})} = \frac{Pr(\mathrm{E/H})}{Pr(\mathrm{E})}$$

tells us that the probability of hypothesis H given the evidence E, *Pr*(H/E), is greater than the prior probability of H, *Pr*(H), in the same proportion. The evidence E raises the probability of hypothesis H; in short, E *confirms* H. Confirmation of a hypothesis by evidence warrants increased credence in the hypothesis, although it might not warrant changing the balance of credence between the hypothesis and an alternative. Evidence may confirm H without being *diagnostic* as between H and an alternative, H´.

15.4.2 Jumping to conclusions and attributional style

These explanatory processes of hypothesis generation and confirmation might, in principle, depart from normality although Maher says (1999, p. 550): 'The processes by which deluded persons reason from experience to belief are not significantly different from the processes by which non-deluded persons do.' Stone and Young (1997) note that normal belief formation is 'fallible' and 'subject to various biases' (1997, p. 332). They argue that a complete answer to the question where a delusion came from will need to appeal to 'a theory of the reasoning biases that lead to the delusional interpretation of the [anomalous experience]' (*ibid.*, p. 341).

As examples of these reasoning biases, Stone and Young mention the *jumping to conclusions* (JTC) bias studied by Garety and colleagues (for reviews, see Garety and Freeman, 1999; Fine *et al.*, 2007) and biased *attributional style*, particularly the

[3] The word 'doxastic' means pertaining to belief or opinion.

externalizing attributional style that seems to play a role in persecutory delusions (for reviews see Bentall *et al.*, 2001; Blackwood *et al.*, 2001; but see Freeman, 2007, p. 440, for the view that 'the empirical case for persecutory delusions being associated with an excessive externalizing style for negative events is unconvincing at present'). These are appropriate examples of biases that might be at work as a subject tries to explain an anomalous experience. The subject's attributional style might bias the generation and consideration of an explanatory hypothesis and the JTC bias might then lead the subject to consider a smaller-than-normal amount of evidence before regarding the hypothesis as adequately confirmed and proceeding to adopt it as a belief.

Stone and Young describe 'the reasoning style of people [with] delusions' as 'the second factor' (1997, p. 346) and they conceptualize this second factor in terms of biased resolution of a permanent tension in the processes of belief formation. It is important to note, however, that what Stone and Young call the second factor is conceived as playing a rather different role from the second factor in the two-factor framework. They are primarily concerned with the processing stage that leads from experience to belief so that their second factor provides part of the answer to the question where the delusion came from. It corresponds to the second stage in the two-factor/three-stage framework.

15.4.3 Observational adequacy, explanatory adequacy, and conservatism

Drawing on Fodor (1987, 1989), Stone and Young propose that there is (1997, p. 349): 'a tension between forming beliefs that require little readjustment to the web of belief (conservatism) and forming beliefs that do justice to the deliverances of one's perceptual systems [beliefs that are observationally adequate]'. In a case of delusion, the balance between these two requirements 'goes too far towards observational adequacy as against conservatism' (*ibid.*).

The idea that a delusion results from a bias towards 'do[ing] justice to the deliverances of one's perceptual systems' is easily appreciated in cases where perceptual experience encodes the content of the delusion and the processing stage that leads from experience to belief involves *endorsement*. Indeed, Stone and Young describe the requirement of observational adequacy as 'seeing is believing' (1997, p. 349). A similar idea also has clear appeal when the anomalous experience is less specific in content and the processing stage that leads from experience to belief involves *explanation*.

Maher proposes that feelings of significance arise from the operation of a comparator or 'detector of changes':

> Survival requires the existence of a detector of changes in the normally regular patterns of environmental stimuli, namely those that are typically dealt with automatically. The detector functions as a general non-specific alarm, a 'significance generator', which then alerts the individual to scan the environment to find out what has changed.
>
> (1999, p. 558)

The normal operation of this device generates feelings of significance in daily life and its pathological operation may give rise to anomalous experiences):

> The origins of anomalous experience may lie in a broad band of neuropsychological anomalies. These include, but are not confined to . . . endogenous neural activation of the

> feeling of significance normally triggered by pre-conscious recognition of changes in a familiar environment.
>
> (*ibid*., p. 551)

From the subject's point of view, a feeling of significance demands explanation in terms of something that has changed.[4] The feeling may be general, occurring in many contexts and accompanying many perceptual experiences. If no change can be detected that would explain the persistent feeling of significance, then an apocalyptic hypothesis might be generated and considered. 'Everything must have changed in some fundamental way' (Maher, 1999, p. 560); perhaps the end of the world is coming (Arthur, 1964, p. 106). The feeling of significance may, however, attach only to particular experiences, such as the subject's experience of an arm (in fact, the subject's paralysed left arm) or of a person (in fact, the subject's spouse). The subject's experience is suffused with a feeling of significance and cries out for explanation in terms of change in the object, person, or situation perceived.

Whether the feeling of significance is general or more particular, trying to do justice to such an experience by postulating change will, inevitably, require adjustment to the preexisting web of belief. Explaining the experience in terms of global change – 'The end of the world is coming' – or in terms of local change in an arm or a person – 'This is not my arm', 'My wife has been replaced by an impostor' – is liable to take the subject far from the requirements of conservatism.

Stone and Young speak of a balance between 'two imperatives' of observational adequacy and conservatism (1997, p. 349). With the distinction between endorsement and explanation in place, we propose to add a third imperative of explanatory adequacy. The imperatives of observational adequacy and explanatory adequacy may both be in tension with the imperative of conservatism, which corresponds to the inertia exerted by a preexisting web of belief. The imperative of observational adequacy corresponds to the prepotent doxastic response of treating a perceptual experience as veridical (seeing is believing). The imperative of explanatory adequacy corresponds to a prepotent doxastic tendency towards acceptance of a hypothesis that explains a salient piece of evidence and is thereby confirmed.[5]

---

[4] Kapur (2003, 2004; Kapur *et al*., 2005) proposes that, in schizophrenia, delusions arise as the patient attempts to make sense of experiences of 'aberrant salience' that result from dysregulated dopamine transmission. For discussion, see Broome and colleagues (Broome *et al*., 2005b).

[5] For an earlier discussion of two ways of interpreting Stone and Young's (1997) suggestion about observational adequacy and conservatism, see Davies and Coltheart (2000, pp. 18–20). The imperative of explanatory adequacy might be conceived as an aspect of a 'theory drive' (Gopnik, 1998, p. 101): 'a motivational system that impels us to interpret new evidence in terms of existing theories and change our theories in the light of new evidence.'

In a study of orientation to uncertainty, Schuurmans-Stekhoven and Smithson (submitted) investigate two dispositions, need for discovery (a tendency to up-date existing beliefs) and need for certainty (a tendency to maintain incumbent beliefs). They show that it is quite possible for someone to have both inclinations to a strong degree, but unlikely that they will lack both of them, and they suggest that the scales may predict biases in belief formation and perhaps the onset or maintenance of delusions.

15.4.4 Parametric variation

The question whether any actual cases of delusion fit the endorsement, rather than the explanationist, model is contested. Fine, Craigie, and Gold (2005) raise problems for both styles of account of the Capgras delusion. Coltheart (2005) defends the explanationist account and says that the endorsement account 'requires much more fleshing out before it will be possible to decide whether it is a viable competitor to the "explanation" account' (2005, p. 153). In contrast, Bayne and Pacherie 'prefer the endorsement version' (2004, p. 4). Jeannerod and Pacherie (2004) provided a detailed account of experiences of agency that would complement an endorsement account of delusions of control in individuals with schizophrenia. Hohwy and Rosenberg (2005) offer an account of the alien control delusion that begins from the hypothesis that 'delusions arise when unusual experiences are taken as veridical' (2005, p. 144).

It seems likely that some cases of delusion will fit the endorsement model and others the explanationist model. If that is right, then the nature of the processing stage that leads from anomalous experience to delusional belief will be one locus of parametric variation within the two-factor framework for explaining pathologies of belief. In fact, the setting of the endorsement/explanation parameter is likely to be a matter of degree and when (or to the extent that) a case fits the explanationist model, there may be further variation amongst accounts of normal, biased, or impaired hypothesis generation and confirmation.

If the processing stage that leads from experience to belief is biased or impaired, then it might, in principle, yield a delusional belief by flawed explanation, or misinterpretation, of quite ordinary or perhaps ambiguous – but not anomalous – experiences. Since the two-factor framework is also a three-stage framework, it can allow for the possibility that there might be no departure from normality earlier than the second stage. This option for parametric variation within the two-factor framework may be relevant to the explanation of some persecutory delusions (Bentall *et al*., 2001; Blackwood *et al*., 2001; Freeman, 2007).

We have been assuming that the first deficit gives rise to an anomalous experience from which personal-level processes of endorsement or explanation lead to belief. But the neuropsychological version of the two-factor framework is officially neutral on the question whether the first deficit gives rise to an anomalous conscious experience. It may be that personal-level processes have no role to play and that the route from first deficit to belief lies wholly at the sub-personal level and involves wholly unconscious processes (Coltheart, 2007, p. 1044, footnote 4). This is a further example of parametric variation that is allowed by the two-factor framework. The route from first deficit to belief might lie mainly at the personal level or mainly at the sub-personal level. If the bottom-up psychological processes that lead to belief are opaque to the subject, then it seems likely that the belief will be pathological in the sense that we mentioned near the outset. The subject will be beset by a belief for which he or she can offer no grounds.

15.5 The second factor: why does the patient not reject the belief?

In any case of delusion, even when we have answered the first question – Where did the delusion come from? – there is a second question: Why does the patient not reject the belief? Suppose a patient has adopted a false proposition ('This is not my arm' or

'My wife has been replaced by an impostor') as a belief. Suppose the answer to the first question appeals to the patient's endorsement or explanation of an anomalous experience. The patient's initial adoption of the belief was a manifestation of a prepotent doxastic response to a perceptual experience or of a prepotent tendency towards acceptance of a confirmed hypothesis. Still, why does the patient not subsequently reject the belief on the grounds of its implausibility and its incompatibility with a mass of available evidence?

According to the two-factor framework, the answer to this question is that the patient has an impairment of belief evaluation. Coltheart proposes that the impairment 'is the same in all people with monothematic delusion' (2005, p. 154) although the impairment is 'very poorly specified' (*ibid.*). We do not yet have an account of the cognitive nature of the second factor in the two-factor framework. To say that the patient does not reject the belief because he or she has lost the ability to make appropriate use of evidence and plausibility in evaluating and revising beliefs (Davies *et al.*, 2001, p. 149) scarcely goes beyond reiterating the fact that the patient's belief is a delusion.

Although the second factor is poorly specified in terms of cognitive function, there are some suggestions that it is a neuropsychological deficit whose neural basis lies in damage to the right hemisphere. Coltheart describes the second deficit in this way.

> The patient has right-hemisphere damage (i.e., damage to the putative belief evaluation system located in that hemisphere).
>
> (2007, p. 1047)

He goes on to review evidence that 'it is specifically *frontal* right-hemisphere damage that is the neural correlate of the impairment of belief evaluation' *(ibid.*, p. 1052).

15.5.1 Evidence and implausibility

Suppose (as before, section 15.4.1) that the evidence, E, provided by an anomalous experience confirms an explanatory hypothesis, H, which is initially adopted as a belief in response to the imperative of explanatory adequacy. In principle, the explanatory hypothesis may be subsequently evaluated in at least two ways. First, the support that the evidence E provides for H is defeasible. Although E confirms H, the totality of the available evidence, including E, may disconfirm H. Second, a hypothesis H that is confirmed by evidence E, and even by the totality of the available evidence, may still have a relatively low posterior probability if it has a very low prior probability.

Posterior probability depends on both degree of (dis)confirmation and prior probability. So the case for rejecting a hypothesis may sometimes depend primarily on the weight of disconfirming evidence and sometimes on the low prior probability of the hypothesis being true. The same grounds for rejection – evidence or implausibility – may apply to a false proposition that is initially adopted as a belief by way of endorsement, rather than explanation, of an anomalous experience.

Later in this chapter (section 15.9.3), we shall arrive at a proposal about the cognitive nature of the second factor that is somewhat informative but also suitably general. Nevertheless, we should be open to the possibility that the answer to the question why the patient does not reject the belief may vary in its details. Some patients may fail to reject their false belief because they do not make proper use of available disconfirming evidence, others because they do not take proper account of the belief's implausibility. This may be another locus of parametric variation within the two-factor framework.

In the Capgras delusion, the most obviously available evidence – the appearance of the patient's wife and her own statements – does not disconfirm the patient's false belief. After all, a good impostor would look like the patient's wife and would say that she was the patient's wife. The evidence confirms both the true hypothesis (that the person is the patient's wife) and the impostor hypothesis, but is not diagnostic as between them. The impostor hypothesis might be regarded as similar to sceptical hypotheses (such as Descartes's evil demon hypothesis) in being 'unfalsifiable'. What counts against the impostor hypothesis is primarily the fact that it is implausible, not only in the view of people without delusions, but also in the light of the patient's other beliefs.

In somatoparaphrenia, a patient's denial of ownership of the left hand may go against available evidence. Bisiach and Geminiani describe the case of patient LA-O:

> On request, she admitted without hesitation that her left shoulder was part of her body and *inferentially* came to the same conclusion as regards her left arm and elbow, given, as she remarked, the evident continuity of those members. She was elusive about the forearm but insisted on denying ownership of the left hand. . . . She could not explain why her rings happened to be worn by the fingers of the alien hand.
>
> (1991, pp. 32–33)

Here, the presence of LA-O's own rings on the fingers of the hand confirms the hypothesis that the hand is hers and disconfirms her belief that the hand is alien.[6]

15.5.2 Subverting the role of evidence and implausibility

In some cases, a patient's initial adoption of a false belief subverts the disconfirmatory role of evidence. Young and Leafhead (1996) describe the case of a 29-year-old woman, JK, who claimed that she was dead (Cotard delusion; Cotard, 1882). They investigated whether patient JK regarded the fact that she had thoughts and feelings as evidence against her belief that she was dead:

> We therefore asked her, during the period when she claimed to be dead, whether she could feel her heart beat, whether she could feel hot or cold, and whether she could feel when her bladder was full. She said she could. We suggested that such feelings surely represented evidence that she was not dead, but alive. JK said that since she had such feelings even though she was dead, they clearly did not represent evidence that she was alive.
>
> (1996, p. 158)

Patient JK accepted that, in general, the probability that someone would have thoughts and feelings while dead was low. But she was convinced that she herself was dead and she regarded her own situation – a dead person experiencing bodily sensations – as unique. In contrast, McKay and Cipolotti (2007) present a case of the Cotard delusion in which evidence did play a disconfirmatory role.

A patient's initial adoption of an explanatory hypothesis as a belief may also subvert arguments for rejecting the hypothesis on the grounds of its implausibility. In somatoparaphrenia, a patient may deny ownership of his left hand and claim that it belongs to someone else. Patient PR (Bisiach, 1988) claimed that his left hand belonged to the

[6] Similar evidence might be presented to a patient with a misidentification delusion (Breen *et al.*, 2002).

examiner and that the examiner had three hands – an implausible view. In an oft-quoted exchange, the examiner highlighted the implausibility of patient PR's belief, asking, 'Ever see a man with *three* hands?' (Bisiach, 1988, p. 469). The patient replied (*ibid.*): 'A hand is the extremity of an arm. Since you have three arms it follows that you must have three hands.' Patient PR believed that an arm – in fact, his own left arm – was not his. From his point of view, it was more plausible that the arm belonged to someone else, such as the examiner, than that it belonged to him. Given that starting point, it was not especially implausible that the examiner should have three hands.

15.6 The task of belief evaluation

In response to the power of a hypothesis to explain an anomalous experience, a patient may accept the hypothesis and regard competing hypotheses as correspondingly improbable. But, normatively speaking, the patient's acceptance of the hypothesis may be unwarranted. In a theoretical paper, Hemsley and Garety suggest (1986, p. 52): 'A normative theory of how people *should* evaluate evidence relevant to their beliefs can provide a conceptual framework for a consideration of how they do *in fact* evaluate it.' In the spirit of Hemsley and Garety's suggestion, we consider the task of belief evaluation in the light of the normative standards of probability theory.[7]

15.6.1 Alternative explanatory hypotheses

Suppose that hypothesis H adequately explains a patient's anomalous experience in the following sense. If E is the evidence provided by the experience then the probability of E given H, *Pr*(E/H), is close to 1, and is higher than the prior probability of E, *Pr*(E). According to Bayes's theorem, the evidence confirms the hypothesis. It raises the probability of the hypothesis in the same proportion as the hypothesis raises the probability of the evidence. But if the prior probability of H is very low then the posterior probability of H may still be low.

We have conjectured that, corresponding to the imperative of explanatory adequacy, there is a prepotent doxastic tendency towards acceptance of a confirmed hypothesis (section 15.4.3). But a patient who accepts H just because it is confirmed by evidence that it explains may, in effect, be underestimating the probability of that evidence given the negation of the hypothesis. The patient may be ignoring an alternative hypothesis, H′, inconsistent with H, that has a higher prior probability than H and is no less adequate to explain the evidence.

Here, we should consider Bayes's theorem in the form:

$$\frac{Pr(\mathrm{H/E})}{Pr(\mathrm{H'/E})} = \frac{Pr(\mathrm{E/H})}{Pr(\mathrm{E/H'})} \cdot \frac{Pr(\mathrm{H})}{Pr(\mathrm{H'})}$$

Suppose that two competing hypotheses, H and H′, are equally adequate to explain the evidence E; that is, suppose that $Pr(\mathrm{E/H}) = Pr(\mathrm{E/H'})$. Then the posterior probabilities of the hypotheses stand in the same ratio as the prior probabilities. The evidence is not

[7] In pursuing this strategy it is, of course, important not to lose sight of the distinction between the normative and the descriptive (see Stone and Young, 1997, p. 342).

diagnostic; it does not change the balance of probabilities between the competing hypotheses. Suppose, for example, that the prior probability of hypothesis H′ is ten times that of H. Then the posterior probability of H′ is also ten times that of H and so the posterior probability of H, *Pr*(H/E), must be less than 0.091.

Hemsley and Garety describe a case of this kind:

> For example, one patient took the appearance of a police car in a busy thoroughfare as unequivocal evidence that the police were chasing him, neglecting the probability of this event occurring if the police had no interest in him.
>
> (1986, p. 53)

The patient's hypothesis that the police were chasing him was, let us agree, adequate to explain the evidence of the police car's appearing on the street. But that evidence cannot shift the balance of probabilities in favour of the patient's hypothesis and against an alternative hypothesis if the alternative hypothesis is also explanatorily adequate. A suitable alternative hypothesis, with a higher prior probability than the patient's hypothesis, would be that the police were not chasing the patient but were chasing someone else in the area.

15.6.2 Alterative explanations, jumping to conclusions, and the confirmation bias

In common parlance, a subject who accepts a hypothesis just because it is adequate to explain a piece of evidence might be described as jumping to a conclusion. So it is important to consider the relationship between the phenomenon that we have been describing (in section 15.6.1) and the JTC bias studied by Garety and colleagues using the beads task (Huq *et al.*, 1988; Garety *et al.*, 1991). In the beads task, subjects are presented with two jars, one jar (A) containing (for example) eighty-five black beads and fifteen yellow beads and the other jar (B) containing eighty-five yellow and fifteen black beads. Subjects are told that initially each jar is equally likely to be chosen, that one will be chosen, and that beads will then be drawn, sequentially and with replacement, from the chosen jar. The subject's task is to decide whether the experimenter is drawing beads from jar A or from jar B. The typical finding is that, by comparison with clinical and non-clinical control participants, patients with delusions ask for fewer beads to be drawn before they reach a decision (which is usually correct). The JTC bias has also been found in individuals at high risk for psychosis (Broome *et al.*, 2007), in relatives of patients with psychosis (Van Dael *et al.*, 2006), and in delusion-prone members of the general population (Linney *et al.,* 1998; Colbert and Peters, 2002).

Freeman and colleagues (Freeman *et al.*, 2004) investigated whether patients with delusions were able to suggest any alternative explanation for their experiences, even if they thought the alternative very unlikely. They also assessed the JTC bias in these patients. About a quarter of the patients were able to suggest an alternative explanation and these patients showed a lesser JTC bias than those offering no alternative explanation. Freeman and colleagues suggest a causal connection between the JTC bias and failure to consider alternative explanations: 'Rapid acceptance of judgments is likely to limit consideration of alternative explanations' (2004, p. 672); 'It is plausible that a more cautious reasoning style may tend toward consideration of alternatives' (*ibid*, p. 678). Indeed, in the literature on

delusions, the phenomenon of ignoring alternative explanations and the JTC bias are often presented as being closely linked (e.g. Stone and Young, 1997, p. 341). They seem, however, to be conceptually distinct.

The phenomenon that we have been describing involves three important features. First, the subject accepts a hypothesis with a relatively low posterior probability, *Pr*(H/E). Second, the subject underestimates the probability of the evidence given the negation of the hypothesis, *Pr*(E/not-H). Third, the subject ignores alternative explanations. In contrast, subjects in the beads task do not accept a hypothesis with a low posterior probability. The probability that jar A has been chosen is 0.85 given that the first bead presented is black and 0.97 given that the first two beads are black. Also, there is very little evidence that the JTC bias involves subjects underestimating the probability of the presented evidence given the negation of the favoured hypothesis. Furthermore, since it is explicit that only two hypotheses are relevant in the beads task, there is no possibility that subjects could ignore alternative explanations of the evidence presented to them. The nature of the connection between failure to consider alternative explanations and the JTC bias requires further theoretical and empirical investigation.

Failure to consider alternative explanatory hypotheses seems to be related to the confirmation bias (Wason, 1960; for a review, see Nickerson, 1998). In an interesting pilot study, Freeman and colleagues (Freeman *et al*., 2005) used Wason's (1960) 2–4–6 task to assess confirmatory reasoning in non-clinical individuals. In this task, participants are told that the experimenter has in mind a rule that classifies ordered triples of numbers. Participants are told that the triple 2–4–6 conforms to the rule and are asked to try to discover the rule by suggesting additional triples for which feedback will be provided. (Participants are told whether or not their suggested triple conforms to the rule.) The triple 2–4–6 suggests the rule 'successive even numbers' and the typical finding is that participants suggest many triples that conform to that hypothesized rule, such as 6–8–10 or 20–22–24. The feedback confirms their initial hypothesis and, because they do not try out triples that are inconsistent with the 'successive even numbers' rule (such as 3–5–7 or 1–2–3), participants may not discover that the actual rule is 'any three numbers in ascending order'.

In the study by Freeman and colleagues (2005), participants who suggested only triples that conformed to the rule that they (at that time) considered likely to be correct were said to show a confirmatory reasoning style. Participants who sometimes suggested triples that did not conform to the rule that they considered likely to be correct were said to show a disconfirmatory reasoning style. Intellectual and executive functioning, psychological symptoms, and delusional ideation were also assessed and participants completed a belief-evaluation task modelled on cognitive therapy.

The findings of the study were that individuals who adopted a disconfirmatory reasoning style in Wason's 2–4–6 task gathered more evidence before reaching a decision, and considered a greater number of hypotheses, than individuals with a confirmatory reasoning style. They had higher IQ scores and lower depression scores and, in the belief-evaluation task, they produced more evidence, both for and against their beliefs. Similar investigations of patient populations would appear to hold considerable promise for both theoretical understanding and therapeutic intervention.

15.7 Acceptance and subversion

A patient's acceptance of hypothesis H on the basis of evidence E is normatively not warranted if a competing hypothesis, H′, also explains evidence E and has a higher prior probability than H. But, once the hypothesis H has been accepted, the power of the competing hypothesis to explain the evidence cannot shift the probabilities in favour of H′ and against H since H is also explanatorily adequate. Having accepted H, the patient assigns a high probability to it and a correspondingly low probability to competing hypotheses.

An argument for rejecting an explanatory hypothesis on the grounds of its implausibility needs to be deployed while the patient still regards the hypothesis as somewhat improbable and when a more probable, and no less explanatory, hypothesis is also available. Thus, for example, there are potential benefits in providing clinical intervention for individuals identified as being in the 'at risk mental state' but before the first episode of psychosis (Broome *et al.*, 2005a). A patient's unwarranted acceptance of a hypothesis is apt to subvert an argument for rejecting it on the grounds of its implausibility, even if an alternative explanation is presented. In order to make proper use of considerations of implausibility when evaluating explanatory hypotheses, the patient needs to take a step back from his or her initial acceptance. The patient must, at least suppositionally, regard the question of the truth or falsity of the hypothesis as open. The patient must then attempt to settle the question whether the hypothesis is true or false by evaluating it alongside alternative explanatory hypotheses.

In his influential book on inference to the best explanation, Lipton (2004) describes the two-stage process of hypothesis generation and selection. First, a shortlist of explanatory hypotheses is generated; second, the best candidate on the shortlist is selected. When the mechanisms of hypothesis generation work well they favour 'those that are extensions of explanations already accepted' (2004, p. 151) and Lipton suggests that this may explain our normal conservatism (*ibid.*): 'Our method of generating candidate hypotheses is skewed so as to favor those that cohere with our background beliefs, and to disfavor those that, if accepted, would require us to reject much of the background.' In some cases of delusion, background beliefs do not adequately constrain hypothesis generation and, as a result, the best candidate on the shortlist is not good enough. This unsatisfactory situation is particularly difficult to rectify if the selection process is allowed to go ahead – if the best candidate on the shortlist is selected and is, so to speak, installed in the advertized position. That is, flawed hypothesis generation is difficult to rectify if, in accordance with a prepotent doxastic tendency, one hypothesis from the inadequately constrained shortlist is selected and adopted as a belief.[8]

We have been considering the task of belief evaluation from a theoretical perspective. On that basis, we can put forward an initial suggestion about cognitive processes that may be implicated in belief evaluation. First, the evaluation of competing hypotheses in the light of evidence and plausibility will involve working-memory resources for the maintenance and manipulation of information. Second, if there is a prepotent doxastic

[8] We are indebted to Tony Stone for drawing our attention to Lipton's (2004) discussion of hypothesis generation in normal inference to the best explanation.

tendency towards accepting a hypothesis that explains a salient piece of evidence and is thereby confirmed, then the step back from initial acceptance will involve executive processes of inhibition.

We now describe putative processes of belief formation and belief evaluation in the Capgras delusion, beginning from the assumption that an anomalous experience figures in the answer to the question where the delusion came from.

15.7.1 Considering alternatives in the Capgras delusion: the explanationist account

Ellis and Young (1990) propose that the anomalous experience in the Capgras delusion arises from disruption of the connection between the patient's face-recognition system and autonomic nervous system. In a development of this proposal, Ellis and Lewis (2001) suggest that '[an integrative device] would . . . compare the expected affective response [i.e. expected on the basis of the activity in the primary face-recognition system] with the actual affective response and some kind of attribution process would take place' (2001, p. 154). Coltheart makes a similar suggestion in terms of unconscious processes of prediction and comparison (2005, p. 155): 'the unconscious system predict[s] that when the wife is next seen a high autonomic response will occur, detect[s] that this does not occur, and report[s] to consciousness, "There's something odd about this woman"'.

These suggestions and Maher's (1999) proposals converge on the idea that, as the result of a neuropsychological deficit and the subsequent operation of a comparator system, the Capgras patient has an anomalous experience.[9] It is a perceptual experience of his wife that is suffused with a feeling of heightened significance, an experience that cries out for explanation in terms of change, or 'something odd', in the immediate environment and particularly in the woman perceived. The hypothesis, H, that the woman perceived is not really the patient's wife seems adequate to explain the patient's anomalous experience. The negation of that hypothesis – that is, the hypothesis, not-H, that the woman perceived is *not* an impostor and really *is* the patient's wife – does not seem to offer the same explanatory promise.

The patient underestimates the probability of the evidence given the hypothesis not-H and ignores – does not even consider – the alternative, more specific (and correct) explanatory hypothesis, H′, that the woman is his wife *and he has suffered a brain injury*. The hypothesis H is, to some degree, confirmed by the evidence and, as the result of a prepotent tendency, is adopted as a belief. This woman, who looks like the patient's wife and says that she is the patient's wife, is not really his wife; she is an impostor.

If this initially adopted belief is to be evaluated and ultimately rejected, then the patient must step back from his acceptance of the impostor hypothesis. But stepping back is not

[9] Here, we envisage abnormally reduced autonomic activity and normal operation of the comparator. Alternatively, an anomalous experience might be produced by abnormal operation of the comparator itself or perhaps by an abnormality that is causally downstream from the comparator (Maher, 1999, p. 551).

sufficient by itself since the false belief is liable to be reinstated by renewed demands for explanation each time the patient looks at his wife. The patient also needs to undertake an evaluation of the impostor hypothesis and alternative explanatory hypotheses, including the brain injury hypothesis, in terms of their plausibility and in the light of available evidence. The patient needs to inhibit a prepotent doxastic tendency and to undertake a cognitive task that is demanding of working-memory resources.

15.7.2 Considering alternatives in the Capgras delusion: the endorsement account

The same key features are also present according to the endorsement account. We now assume that the representational content of the Capgras patient's experience is more specific than 'This is someone who looks just like my wife *but there is something odd about her*.' It is, rather: 'This is someone who looks just like my wife *but it is not really her*.' The processes that determine this content of experience are not yet specified – the endorsement account, like much else in this area, needs 'fleshing out' (Coltheart, 2005, p. 153). We might conjecture that the generation of the content will – like the attribution process in Ellis and Lewis's (2001) account – involve a comparator system or integrative device that has access to both the primary face-recognition system and the autonomic nervous system.

The prepotent doxastic response to a perceptual experience with this content is to believe that this person, who looks just like the patient's wife, is not really her. This belief is subsequently elaborated into the belief that the patient's wife has been replaced by an impostor. We can conceive of this latter belief as an explanatory hypothesis that the patient has adopted. But it is important to recognize that what is being explained is the fact, as the patient believes, that the person who looks like his wife is not really her:

> The prior possibility that the spouse is an impostor of some sort is of course very low, but if this hypothesis best explains that (as the patient believes is true) the spouse is really a stranger that looks like the spouse, then it is very probable that the spouse is an impostor.
>
> (Hohwy and Rosenberg, 2005, p. 154)

As Hohwy and Rosenberg point out, what needs explaining, from the patient's point of view, is not adequately explained by the hypothesis that the woman is the patient's wife and the patient has suffered a brain injury (*ibid.*, p. 155): 'The brain pathology hypothesis would only be relevant if the patient could accept that what needs explaining is the mere *experience* that it is as if the spouse looks like a stranger.'

Thus, as before, in order to consider and evaluate alternative explanatory hypotheses, the patient first needs to take a step back. The patient must inhibit the prepotent doxastic response of treating a perceptual experience as veridical and, instead, treat the experience as standing in need of explanation.[10]

[10] We note that Hohwy and Rosenberg propose that 'it is intra- and inter-modal reality testing that can inhibit the pre-potent doxastic response to believe what we experience, and when such reality testing procedures are exhausted, nothing else will on its own inhibit the pre-potent response' (2005, p. 149). They also suggest that 'unusual beliefs arise when unusual

15.8 Interlude: hypotheses, beliefs, and evaluation

We have interpreted the question, 'Where did the delusion come from?', as asking how the patient came to adopt the false belief and we have considered belief evaluation under the assumption that, normatively, it is primarily supposed to take place after initial adoption of a belief. Thus, we have interpreted the question, 'Why does the patient not reject the belief?' as asking why the patient is not able to evaluate and reject the belief that has been initially adopted.

An alternative approach would be to interpret the first question as asking how the patient came to entertain or consider the false hypothesis. This may be what Coltheart intends when he glosses the first question as (2007, p. 1047): 'what is responsible for the *content* of the particular belief?' According to this approach, belief evaluation – really, hypothesis evaluation – begins earlier and the second question asks, in part, why the patient is not able to evaluate the hypothesis and reject it, instead of adopting it as a belief. (Since a delusion is a false belief that is maintained, the second question must also ask why the patient is not able to evaluate and reject the belief even after adopting it.)

The difference between these approaches seems to correspond to a difference between two accounts of normal believing. In a series of papers, Gilbert and colleagues (Gilbert *et al.*, 1990, 1993; Gilbert, 1991) have contrasted Cartesian and Spinozan views of belief and have presented experimental results in support of the Spinozan view.[11] Each view of belief can be summarized in terms of two stages, a representation stage and an assessment stage. On the Cartesian view, the representation stage involves *comprehension*, which 'precedes and is separate from assessment' (Gilbert, 1991, p. 108). A hypothesis is grasped and then, in the assessment stage, the hypothesis is either *accepted* as true and adopted as a belief, or else *rejected* as false. On the Spinozan view, in contrast, the representation stage involves both comprehension and *acceptance* (*ibid.*, p. 107): 'People believe in the ideas they comprehend, as quickly and automatically as they believe in the objects they see.' Then, in the assessment stage, the already adopted belief is either *certified* or else *unaccepted*.

Our approach in this chapter is influenced by the Spinozan view of belief as Gilbert presents it but recent experimental findings suggest that 'The relation between

experiences are taken as veridical because they occur in sensory modalities or at processing stages where application of the available reality testing procedures keeps giving the same result and where further intra- or inter-modal reality testing cannot be performed' (*ibid.*, p. 153). Their overall position is that in cases of delusion, such as the Capgras delusion or the alien control delusion, further reality testing is not possible. Consequently, they regard the transition from anomalous experience to delusional belief as unavoidable (p. 156) and reject the basic argument for a two-factor framework. That is, they reject the claim that there are patients who have the first factor that is implicated in a delusion – a particular kind of anomalous experience arising from a neuropsychological deficit – yet do not have the delusion.

[11] We are grateful to Tony Stone for many conversations about the work of Gilbert and colleagues and we acknowledge the influence of his presentation, 'Delusions: Learning from Spinoza', at a workshop on delusion and self-deception held at Macquarie University in November 2004. We shall not engage with questions about the relationship between the two views that Gilbert contrasts and the historical philosophers for whom they are named.

comprehension and belief is a complex one' (Hasson *et al.*, 2005, p. 571). We acknowledge that the Cartesian view, or some hybrid, may provide a better account of normal believing. This acknowledgement might seem to pose a threat to our suggestion that belief evaluation involves inhibitory executive processes. If the processes of evaluation begin before any belief is adopted, then no 'step back' from initial acceptance of a hypothesis is required.

In fact, however, the suggestion is not threatened since the influence of prepotent responses and tendencies still needs to be inhibited; imperatives still need to be resisted. As Gilbert, Tafarodi, and Malone say, in a discussion of sceptical doubt:

> For Descartes, being skeptical meant understanding an idea but not taking the second step of believing it unless evidence justified taking that step. For Spinoza, being skeptical meant taking a second step backward (unbelieving) to correct for the uncontrollable tendency to take a first step forward (believing). Both philosophers realized that achieving true beliefs required that one subvert [inhibit, resist] the natural inclinations [prepotent tendencies, imperatives] of one's own mind; for Descartes this subversion was proactive, whereas for Spinoza it was retroactive.
>
> (Gilbert *et al.*, 1993, p. 230)

Stone and Young (1997) say that, in cases of delusion, the balance between imperatives goes too far towards observational adequacy – or, we have added, explanatory adequacy – and departs too far from conservatism. If this is where a delusion came from then, it may seem, evaluating and rejecting the belief involves inhibiting the natural inclinations towards observational or explanatory adequacy and restoring the influence of the imperative of conservatism. But merely allowing the preexisting web of belief to exert inertia, so that an observationally or explanatorily adequate hypothesis is not accepted, is not yet sufficient for belief evaluation. What is required is that the patient should assess competing hypotheses (by weighing evidence and plausibility) while also controlling and balancing (inhibiting or not) the influences of observational adequacy, explanatory adequacy, *and* conservatism. (In the case of anosognosia, the belief that needs to be rejected is part of the patient's preexisting web of belief.) Cognitive tasks with this structure – undertaking an analytic assessment while controlling heuristic influences – are the focus of dual-process accounts of reasoning.

15.9 **Dual-process accounts of reasoning**

Dual-process accounts propose that there are two quite different kinds of cognitive processes involved in reasoning – and also in judgement and decision-making (Evans, 2003). The two kinds of processes are sometimes referred to as 'System 1' versus 'System 2' processes although, as Evans (2007) says, 'the mapping of dual processes on to underlying dual systems is fraught with difficulties' (2007, p. 322) System 1 or *heuristic* processes are 'rapid, preconscious, and computationally powerful'; System 2 or *analytic* processes, in contrast, are 'slow, sequential, and effortful' (*ibid.*; see also Stanovich, 1999). System 1 processes underpin cognitive biases and are heterogeneous in their nature. Some may be evolutionarily ancient, but not all are. System 2

processes permit 'abstract hypothetical thinking that cannot be achieved by system 1' (Evans, 2003, p. 454). A central idea in dual-process accounts is that the two kinds of processes can come into conflict or competition.

In this section, our aim is to draw on research that is motivated by dual-process accounts in order to generate proposals about the cognitive nature and neural basis of belief evaluation. We shall connect the two areas by considering the imperatives of observational and explanatory adequacy and, particularly, conservatism as belonging with System 1 or heuristic processes and considering the assessment of competing hypotheses as belonging with System 2 or analytic processes. In a similar spirit, Freeman and colleagues suggest that 'belief evaluation may be partly understood by drawing upon the reasoning literature' (Freeman *et al.*, 2005, p. 243). They propose that cognitive therapy for clinical disorders, including the technique of 'encourag[ing] patients to evaluate their beliefs', may 'promote [System 2] analytic reasoning to modify particular conclusions derived from [System 1] processes' (*ibid.*, p. 244).

One important piece of evidence that supports dual-process accounts of reasoning is provided by the *belief-bias* effect (Evans *et al.*, 1983; for a review, see Klauer *et al.*, 2000). The belief bias is 'the tendency for people to judge the validity of an argument on the basis of whether or not they agree with its conclusion' (Evans, 2007, p. 322). Participants are asked to assess syllogistic arguments for logical validity. In some of the arguments, the conclusion is intuitively believable (e.g. Some highly trained dogs are not police dogs); in others, the conclusion is intuitively unbelievable (e.g. Some millionaires are not rich people). Validity of the arguments and believability of their conclusions can be varied independently to generate items of four types: Valid argument–Believable conclusion; Valid argument–Unbelievable conclusion; Invalid argument–Believable conclusion; Invalid argument–Unbelievable conclusion. In the second and third types of argument, there is a *conflict* between the response based on validity and the response based on believability; in the other two types of argument there is *no conflict*.

Participants are explicitly instructed to assume that the premises of the syllogism are true and to judge whether the conclusion necessarily follows from the premises. But the typical finding in these experiments is that participants' responses are influenced by the believability of the conclusion as well as by logical validity. Despite the explicit instructions, it is extremely difficult for healthy adult participants to inhibit the influence of their prior beliefs. The dual-process interpretation of the belief-bias effect is as follows:

> System 2 [analytic] thinking is both volitional and responsive to verbal instructions whereas System 1 [heuristic] thinking is not. Hence System 1 influences – in this case belief bias – can only be suppressed indirectly by asking people to make a strong effort to reason deductively.
>
> (Evans, 2003, p. 456)

As Stanovich (2003) says, there is a 'tendency to automatically bring prior knowledge to bear when solving problems' and this tendency is 'so ubiquitous that it cannot easily be turned off – [it is] a fundamental computational bias' (p. 292).

15.9.1 The role of working memory and inhibitory executive processes

Earlier (section 15.7), we suggested that belief evaluation may involve working memory and inhibitory executive processes. Working-memory tasks are said to involve maintenance and manipulation of information but it is useful to make some distinctions among these tasks. Working-memory capacity is often assessed by span tasks of which the simplest require the subject to reproduce a list of digits or words. An example of such a task is Digit Span Forward, a subtest of the Wechsler Memory Scale–Revised (WMS-R; Wechsler, 1987). The subject is asked to reproduce successively longer lists of digits and the subject's digit span is the length of the longest list that the subject can reproduce correctly. In fact, these simplest span tasks, which involve maintenance but not manipulation of information (storage but not processing), would usually be described as testing attention or short-term memory rather than working memory. One way to introduce manipulation or processing of information is to ask the subject to reproduce a list of digits or words in reverse order, as in the Digit Span Backward subtest of the WMS-R.

In a more complex kind of span task, participants are asked to memorize a list of words while also carrying out simple arithmetical calculations. Each word on the list is preceded by an arithmetical problem, for example:

Is $(4 \div 2) + 3 = 6$? (yes or no) DOG

The number of words that can be recalled provides an estimate of working-memory capacity. This kind of working-memory span task requires both storage (of the words) and processing (for the calculations) and it also requires the participant to maintain task-relevant information (the words) in the face of distraction or interference (from the calculations). That is, the complex span task requires storage and processing of information and also executive processes of controlled attention – using attention to maintain or suppress information (Engle *et al.*, 1999; Engle, 2002). Working-memory capacity as assessed by the complex span task is inextricably linked to executive function and is sometimes referred to as 'executive working memory'. Smith and Kosslyn say (2007, p. 259): 'The central executive is what does the "work" in working memory.'

In an experimental study, De Neys (2006) investigated an assumption of dual-process accounts concerning the role of working memory and inhibitory executive processes in cognitive performance that depends on System 2:

> [T]he two systems [System 1 = heuristic; System 2 = analytic] will sometimes conflict and cue different responses. In these cases, the analytic system will need to override the belief-based response generated by the heuristic system. *The inhibition of the heuristic system and the computations of the analytic system are assumed to draw on executive working memory resources.*
>
> (2006, p. 428; emphasis added)

This assumption yields the prediction that performance on conflict items will be better in participants with higher working-memory capacity and that performance on conflict items will be worse when participants have to perform a secondary task that burdens their executive resources. It is also predicted that neither working-memory capacity nor executive load will affect performance on no-conflict items, since heuristic (System 1) processes will generate the correct response.

In this study, the working-memory capacity of participants was assessed using a complex span task of the kind just described. The primary task was then a syllogistic reasoning task of the kind used in belief-bias experiments and a secondary executive load task required participants to remember a pattern of three (low load) or four (high load) dots in a 3-by-3 matrix. A dot pattern was presented before each syllogism and the participant had to reproduce the pattern after the syllogism had been assessed for logical validity.

This kind of dot-memory task involves maintenance but not manipulation of information and we might not expect that it would impose a load on executive function. Indeed, in the verbal domain, it has been shown that working-memory tasks that involve both storage and processing are more strongly related to executive functioning than tasks that involve storage alone. However, in a study of visuospatial working memory and executive functioning, Miyaki and colleagues (Miyaki *et al.*, 2001) showed that the situation is different in the visuospatial domain. There, both kinds of task, specifically including the dot-memory task, are strongly related to executive functioning.

The main findings of the De Neys (2006) study were these: greater working-memory capacity resulted in better performance on conflict items, while performance on no-conflict items was uniformly high. Executive load had a negative impact on performance on conflict items but did not affect performance on no-conflict items. These findings support the assumption that 'the inhibition of the heuristic system and the computations of the analytic system . . . draw on executive working memory resources' (2006, p. 428). The results of the study show that 'erroneous reasoning in the case of belief-logic conflict is not only associated with, but also directly caused by, limitations in executive resources' (*ibid.*, p. 432).

It is natural to suppose that belief evaluation involves both inhibition of heuristic systems – that is, inhibition of prepotent tendencies or resistance against imperatives – and computations of the analytic system – that is, assessment of competing hypotheses in the light of evidence and plausibility. So we interpret the findings of De Neys's (2006) study as providing some support for the suggestion that belief evaluation involves working-memory resources and inhibitory executive processes. Paraphrasing De Neys, we may also suggest that *erroneous* belief evaluation – maintaining a false belief – in cases where the normative requirements of belief evaluation *conflict* with the imperatives of observational adequacy, explanatory adequacy, or conservatism is *caused by limitations* in executive working-memory resources (that is, working memory and executive function).[12]

15.9.2 The neural basis of performance in a belief-bias experiment

Goel and Dolan (2003) used event-related functional magnetic resonance imaging (fMRI) to investigate the neural basis of performance by subjects in a belief-bias experiment. In particular, they measured neural activation as subjects responded to

[12] It is of some interest to note that, in a study of individuals at high risk for psychosis (Broome *et al.*, 2007), the at-risk group performed significantly worse than healthy control participants on a test of working memory (a simple span task using coloured beads). Also, within the at-risk group, the degree of JTC bias was found to be correlated with the number of errors on the bead span task.

syllogisms in which there was a conflict between the response based on validity and the response based on believability (Valid argument–Unbelievable conclusion and Invalid argument–Believable conclusion). When subjects yielded to the influence of their prior beliefs and gave the logically incorrect response there was activation of ventromedial prefrontal cortex (VMPFC); when subjects inhibited the belief bias and gave the logically correct response there was activation of right inferior prefrontal cortex (2003, pp. B17, B19).[13]

Goel and Dolan say that 'the activation of VMPFC in incorrect trials highlights its role in non-logical, belief-based responses' (p. B19). Here it is of some interest to note the result of a study (Adolphs *et al.*, 1996) using the Wason Selection Task (Wason, 1968). In this task, participants are asked which of four cards they need to turn over in order to decide whether a conditional statement is true or false. When the conditional statement is abstract (e.g. If there is a D on one side of any card then there is a 3 on its other side) very few healthy adult participants (fewer than 10%) make the logically correct response. When the conditional statement is deontic, realistic, and familiar (e.g. If you are in a bar drinking beer then you must be over 18 years old) and participants are asked which cards they need to turn over in order to decide whether anyone is breaking the rule, performance is much better (more than 75% of participants make the logically correct response).

In the study by Adolphs and colleagues (1996), the performance of patients with lesions of dorsolateral prefrontal cortex and control subjects with lesions outside the frontal cortex was facilitated by material that was deontic, realistic, and familiar by comparison with less familiar material. Patients with VMPFC lesions, in contrast, performed no better on the familiar than on the less familiar material. They were unable to make appropriate use of information about familiar situations.

Goel and Dolan also conjecture that 'the right prefrontal cortex involvement in correct response trials is critical in detecting and/or resolving the conflict between belief and logic' (2003, p. B19). One possibility is that this neural activation corresponds to controlled or executive attention that is required to facilitate the performance of the logical task of assessing the validity of the argument in the face of distraction from prior beliefs about the conclusion.

15.9.3 A proposal

We have said that it is natural to suppose that belief evaluation involves System 2 processes including processes of inhibiting the influence of prepotent responses and tendencies. Indeed, belief evaluation seems to be a fine example of the 'abstract hypothetical thinking that cannot be achieved by system 1' (Evans, 2003, p. 454). We now add that Goel and Dolan's (2003) finding of right prefrontal cortex activation when

[13] Goel and Dolan (2003) locate this activation in Brodmann's area 45 (p. B17). This area would be included in dorsolateral prefrontal cortex (DLPFC) on an inclusive use of that term. But, as Goel and Dolan's description 'right inferior prefrontal cortex' indicates, the area would be inferior to right DLPFC on a more restricted use of that term.

subjects give logically correct responses under conditions of conflict seems to be broadly consistent with the suggestion that belief evaluation has a neural basis in the right frontal region of the brain. A consequence of this suggestion would be that 'frontal right-hemisphere damage . . . is the neural correlate of the impairment of belief evaluation [the second factor]' (Coltheart, 2007, p. 1052).[14]

We are now in a position to make a proposal about the cognitive nature and neural basis of the second factor in the explanation of delusions. We can draw on our earlier task analysis (Sections 15.6 and 15.7), Coltheart's (2007) arguments, and the findings from behavioural experiments (De Neys, 2006) and neuroimaging (Goel and Dolan, 2003) using a task that is relevantly similar to belief evaluation. The proposal is that the second factor is an impairment of working memory or executive function with a neural basis in damage to the right frontal region of the brain.

15.10 **Anosognosia as a case study**

Patients with anosognosia for their motor impairments, or for the consequences of their impairments for activities of daily living, have false beliefs that are maintained against the evidence. In one respect, explaining this pathology of belief may be less complex than explaining other delusions, such as somatoparaphrenia or the Capgras delusion. In cases of anosognosia, there is a straightforward answer to the question where the delusion came from. The beliefs that constitute anosognosia are not new and exotic but old and commonplace (though there may seldom have been the occasion to articulate them explicitly). Patients with anosognosia have believed for many decades that they can raise their left arm and move their left leg, or that they can clap their hands and walk upstairs. What is new is that, in the dramatically changed circumstances following a right-hemisphere stroke, these beliefs are no longer true. The question that is pressing is: Why do these patients not reject their false beliefs in the light of the evidence available to them in their changed circumstances?

At this point, we can anticipate two problems. First, the question where the delusion came from has a straightforward answer that does not mention any anomalous experience or neuropsychological deficit. So how can anosognosia fit the two-factor framework for explaining delusions? Second, we have proposed that the second factor in the explanation of delusions is an impairment of working memory or executive processes with a neural basis in damage to the right frontal region of the brain. But how much working memory, and what executive processes, does a stroke patient need in order to recognize that his or her arm is paralysed?[15] We shall address these two problems in turn. But before that, we consider the change in a patient's beliefs when a right-hemisphere stroke causes motor impairments but not anosognosia.

[14] For a review of neuroimaging in individuals with delusions in the context of psychosis, see Broome and McGuire (2008).

[15] John Marshall asked us (personal communication): 'How much WM do I need to notice that my arm is paralysed?'

15.10.1 Motor impairments without anosognosia

Consider a hypothetical case of a patient with left-side paralysis following a right-hemisphere stroke and without anosognosia. The patient may intend and try to raise his left arm, but proprioception and vision will tell him that the arm is hanging by his side. The patient may direct his attention to the left side of his body and confirm that his left arm has not moved. Furthermore, when the patient tries to raise his arm, a comparator within the motor control system will detect a mismatch between the expected movement of the arm and what actually happens and the patient will be alerted to his paralysis (Heilman *et al.*, 1998, p. 1908). In short, the patient will have immediate bodily experiences of his paralysis – experiences with the representational content that his left arm does not move despite his trying. The prepotent doxastic response to experiences with this content is for the patient to believe that his left arm does not move. This may subsequently be elaborated into the belief that the patient's left arm is paralysed.

This belief, adopted in response to the imperative of observational adequacy, is implausible in the light of the patient's preexisting web of belief. But although adopting the belief goes against the imperative of conservatism, other evidence confirms the belief. The patient is unable to lift a glass with his left hand, unable to clap his hands, unable to tie a knot, to shuffle a pack of cards, or to type with both hands. Thus, the patient will be concurrently aware of his motoric failures, will acknowledge his impairments, and will appreciate their consequences for activities of daily living. His long-held but no longer true beliefs will be rejected and his newly adopted beliefs will be maintained.

The patient's change of belief is produced, in the first instance, by bodily experiences of motoric failure. The newly adopted belief is then maintained because of recurrent experiences of paralysis and other confirming evidence. If there were disconfirming, instead of confirming, evidence, then the patient might reject the belief. Suppose the patient were to discover that, despite his bodily experiences of his left arm apparently not moving, he could still lift a glass with his left hand, clap his hands, tie a knot, shuffle cards, and type. Then the patient would quite possibly conclude that he was suffering from anomalous and deeply disconcerting experiences as if he were paralysed but that he was not, in reality, motorically impaired.

15.10.2 Anosognosia in the two-factor framework

A patient who was concurrently aware of his motoric failures might still not acknowledge his motor impairments. This might happen if the patient were to have a memory impairment specific to information about the movements or positions of parts of his body (Carpenter *et al.*, 1995) or if, for some other reason, the information provided in bodily experiences were not consolidated into more lasting representations (House and Hodges, 1988; Marcel *et al.*, 2004). For example, Karnath and colleagues (Karnath *et al.*, 2005) report a neuroimaging study in which anosognosia was found to be associated with damage to the right posterior insula, a structure which 'seems to be involved in integrating input signals related to self-awareness and to one's beliefs about the functioning of contralateral body parts' (2005, p. 7137). Nevertheless, it is

natural to suppose that concurrent *un*awareness of motoric failure may often be a factor in anosognosia for motor impairments. There are several proposals about neuropsychological deficits that would impair a patient's immediate experience of paralysis.

Levine (1990) has proposed that somatosensory loss is a factor in anosognosia and Vuilleumier (2004) that unilateral neglect is 'a notable suspect in anosognosia' (2004, p. 10). Heilman has proposed that paralysis is not detected as the result of an impairment to the intentional-preparatory systems involved in motor control (Heilman, 1991; Heilman *et al.*, 1998). Frith, Blakemore, and Wolpert (2000, pp. 1780–1782) suggest that, because of failure of a comparator within the motor control system, patients may experience illusory movements of their paralysed limbs (see also Feinberg *et al.*, 2000). Patients experience their limbs moving as intended when no movement actually occurs.[16] In patients with one or more of these impairments of sensation, attention, or motor control, long-held beliefs about their left-side limbs may remain somewhat credible.

These impairments are candidate factors in the explanation of anosognosia but they do not provide an adequate answer to the question why these patients maintain, rather than reject, their long-held but now false beliefs. Marcel and colleagues give vivid expression to this inadequacy:

> [I]t is not just that they fail *motorically*. The consequence of such [motoric] failures is that, in trying to get out of bed to go to the toilet or to lift an object, they fall over or incur a similar accident, often lying helpless or hurting themselves.
>
> (2004, p. 35)

Even without the immediate bodily experience of paralysis, patients have a mass of other evidence of their motor impairments. Normatively, they should reject their false beliefs in the light of this evidence.

Furthermore, there are patients whose bodily experience suggests that they can still move their left-side limbs but who nevertheless acknowledge their motor impairments (Marcel *et al.*, 2004). Some patients who have recovered from anosognosia continue to describe bodily experiences of being able to move (Chatterjee and Mennemeier, 1996). One patient, when asked, 'Can you raise the left [arm]?', responded: 'It feels like it's rising, but, it's not' (1996, p. 229). Another patient, HS, reported that the idea that he could move his paralysed limbs still seemed credible even though he was able to reject it:

> E: What was the consequence of the stroke?
>
> HS: The left hand here is dead and the left leg was pretty much.
>
> HS: (later): I still feel as if when I am in a room and I have to get up and go walking . . . I just feel like I should be able to.

[16] Berti and colleagues (Berti *et al.*, 2005) report a neuroimaging study in which anosognosia was found to be associated with damage to 'areas related to the programming of motor acts' (2005, p. 488), including dorsal premotor cortex. They interpret the findings as supporting the hypothesis that there is a degree of commonality in the neural substrates of motor control and awareness of action.

> E: You have a belief that you could actually do that?
>
> HS: I do not have a belief, just the exact opposite. I just have the feeling that sometimes I feel like I can get up and do something and I have to tell myself 'no I can't'.
>
> (*ibid*., p. 227)

We note, however, that the argument for a second factor in anosognosia might not apply in the first few days following a stroke if, because of sedation or lack of arousal, the patient did not try to engage in activities of daily living. In such a case, the evidence that Marcel and colleagues (2004) describe would not be available to the patient and anosognosia for motor impairments would arguably not be a delusion. We should expect that studies of anosognosia in the acute stage following stroke may focus on candidate factors that impair sensation, attention, or motor control. At that stage, neuropsychological deficits causing concurrent unawareness of motoric failure may be sufficient to explain the patient's false belief.

Still, anosognosia beyond the first few days following stroke fits the two-factor framework for explaining delusions. A first factor impairs the patient's concurrent awareness of paralysis. But the first factor is not sufficient; patients with the first factor may not have anosognosia. There must be a second factor that explains why the patient does not make appropriate use of a mass of available evidence. Levine (1990) proposed that, when a patient suffers somatosensory loss, paralysis is not phenomenally immediate. Knowledge of paralysis then requires a process of *discovery*. Anosognosia occurs when the first factor is accompanied by additional impairments that impact negatively on observation and inference so that the patient is 'unable to assimilate information from a variety of sources to form a consistent and accurate judgement' (1990, p. 254). Vuilleumier gives clear expression to a generalization of Levine's proposal (2004, p. 11; emphasis added): '*any* neurological dysfunction susceptible to alter the phenomenal experience of a defect might provide the ground out of which anosognosia can develop when permissive cognitive factors are also present'.

15.10.3 Impaired working memory and executive function in anosognosia

A patient with motor impairments but without impairments of sensation or attention, intentional-preparatory systems or comparator systems, memory or consolidation, would very probably recognize his or her paralysis and would do so relatively immediately, without depending heavily on working memory or executive processing. However, recognition of paralysis is more demanding when, as the result of one or more of these impairments, it is not phenomenally immediate. This is the leading idea of Levine's discovery theory and of our account of anosognosia within the two-factor framework.

It is theoretically useful to consider an analogy with the task of assessing a syllogism for logical validity in a belief-bias experiment. When there is no conflict between the response based on validity and the response based on believability, performance on the task is not affected by the availability of executive working-memory resources because heuristic (System 1) processes are adequate to generate the correct response. When there is a conflict, the task is demanding of executive working-memory resources. If those resources are unavailable – as in individuals with low working-memory capacity or under executive load – then performance on the task is degraded.

In the case of a patient with motor impairments, the fundamental conflict is between the reality of the patient's situation, for which there is a mass of evidence (particularly after the first few days), and the patient's long-held beliefs. If no first factor impairs the patient's concurrent awareness of paralysis, then the imperative of observational adequacy is set against the imperative of conservatism and is liable to lead to revision of those long held beliefs. The conflict between the response based on available evidence and the response based on believability is thus reduced or eliminated. In such a case, the patient's ability to reach an accurate judgement should not depend on the availability of executive working-memory resources. The prepotent doxastic response to the patient's bodily experiences is already adequate to generate a true belief.

When there is a first factor, however, there is a severe conflict between the response based on available evidence and the response based on prior beliefs – and all the more so if illusory limb movements enter the battle on the side of the prior beliefs. In this kind of case, our earlier task analysis and the analogy with belief-bias experiments lead to a prediction. Limitations of executive working-memory resources (that is, working memory and executive function) may well have the consequence that prior beliefs that are now false may be maintained, rather than rejected.

In a study of persisting unilateral neglect, the first author investigated cognitive impairments, including impairments of working memory and executive function, in seven right-hemisphere stroke patients (Aimola, 1999; Maguire and Ogden, 2002).[17] All the patients suffered from unilateral neglect – a 'notable suspect' as a first factor in anosognosia – persisting at least 3 months following their stroke. Four of the seven patients (M1, M3, M5, and M6) demonstrated moderate or severe anosognosia for motor impairments and their consequences for activities of daily living – scores of 2 or 3 on the four-point (0–3) scale proposed by Bisiach (Bisiach *et al.*, 1986; Bisiach and Geminiani, 1991). Two patients (F2 and M2) demonstrated at most mild anosognosia (score of 1)[18] and patient M4 frankly acknowledged his motor impairments and their consequences for everyday activities (score of 0). All the patients showed impairments in at least two of the following four areas: visual or verbal memory (recognition or recall), sustained attention, working memory, and executive function. Impairments of memory and sustained attention were doubly dissociated from anosognosia, but impaired working memory was associated with anosognosia.

One test of working memory was the Elevator Counting with Distraction subtest of the Test of Everyday Attention (TEA; Robertson *et al.*, 1994). This requires the patient to respond to two types of auditory tones by counting the low tones and ignoring the high tones. Three patients (M3, M5, and M6), all with moderate or severe anosognosia, demonstrated problems on this test. The other patient with anosognosia (M1) scored in the normal

[17] Nine patients were studied but two are excluded from the discussion: patient F1, following a left-hemisphere stroke, demonstrated no anosognosia; patient F3 had moderate-to-severe anosognosia but time issues prevented her from completing the full neuropsychological test battery.

[18] We note that some studies use a three-point scale on which patients who acknowledge their impairment in response to a specific question (Bisiach's score of 1) are scored 0 and classified as not having anosognosia (Berti *et al.*, 1996).

range on Elevator Counting with Distraction but demonstrated problems on another test of working memory, the Digit Span Backward subtest of the WMS-R (Wechsler, 1987).

The Elevator Counting with Reversal subtest of the TEA is also proposed to load on the working-memory factor. In this subtest, the patient is presented with tones in three pitches. The high and low tones indicate the direction in which the elevator is travelling, up or down, while the medium tones indicate floors that the elevator passes. The task is to calculate the final position of the elevator by counting the medium tones and using the high and low tones to indicate the direction of counting. Although most of the patients found it much too difficult to keep track of the number of tones while shifting direction of counting (see Robertson *et al.*, 1994), patient F2, with mild anosognosia, and patient M4, with no anosognosia, scored *above* the normal mean.

Executive function was assessed by a computerized version of the Wisconsin Card Sorting Test (WCST; Harris, 1988). The WCST is a demanding test involving several executive functions (Lie *et al.*, 2006) and it is acknowledged that poor performance on the test is difficult to interpret (Cinan and Öktem Tanör, 2002; Lezak *et al.*, 2004). It is, however, suggestive to observe that the only patient to score in the normal range on this test was patient M4, with no anosognosia. The Visual Elevator subtest of the TEA is proposed to load on the same attentional switching factor as the WCST (Robertson *et al.*, 1994). The patient counts successive drawings of elevator doors. Along the way, large arrows pointing either up or down indicate the direction in which counting is to continue. Patient M4 again had the best performance on this task, scoring above the normal mean.[19] We also observe that all of the seven patients except M4 (following a right-basal-ganglia haemorrhage) had lesion locations that included right dorsolateral prefrontal cortex (Fig. 15.1).

In a subsequent statistical analysis,[20] we investigated whether patients' scores on fifteen tests of visuoperceptual function, memory, sustained attention, working memory, and executive function were predicted by their anosognosia scores (0–6, the sum of scores, 0–3, for upper and lower limbs). Only the scores on Elevator Counting with Distraction and the scores on the two measures commonly used to assess the WCST, Categories Achieved and Perseverative Errors, were significantly predicted by the anosognosia scores.

These findings are broadly consistent with the proposal that the second factor in the explanation of delusions, including anosognosia for motor impairments, is an impairment of working memory or executive function with a neural basis in damage to the right frontal region of the brain.

[19] In subsequent work, it will be important to separate out the inhibition of distraction and the attentional switching components of tasks such as Elevator Counting with Distraction and Visual Elevator.

[20] For a brief account of the statistical analysis, see Aimola Davies *et al.* (2009); for a more theoretical discussion, see Smithson *et al.* (submitted).

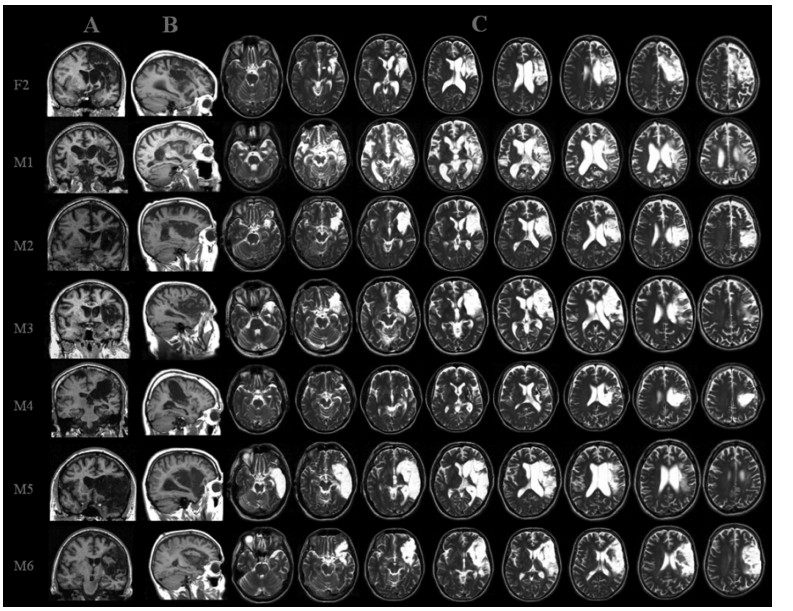

Fig. 15.1 Structural MRI scans: Column A represents a T1-weighted coronal section through the mid-temporal region; Column B represents a T1-weighted sagittal section on the side with the lesion (right in these patients); and Column C displays a series of axial T2-weighted scans in 7.5 mm steps. The right hemisphere is shown on the right. (Thanks to Jerome Maller for preparing the figure.)

15.11 **Summary and conclusion**

In this chapter, we have examined the two-factor framework for explaining pathologies of belief. According to the two-factor framework, we can explain a delusion by answering two questions. First, where did the delusion come from? Second, why does the patient not reject the belief?

We argued that the two-factor framework is also a three-stage framework and that answers to the question where the delusion came from may vary considerably from case to case. We proposed that this heterogeneity could be conceived as parametric variation within a single explanatory framework. There may or may not be abnormality in the first stage. If there is first-stage abnormality then the nature of the abnormality will vary from delusion to delusion and may vary from case to case of the same delusion. A first deficit may or may not give rise to an anomalous experience and the route from first deficit to belief may lie mainly at the personal or the sub-personal level. At the personal level, an anomalous experience may have a representational content close to or far from the content of the delusion itself and the route from experience to belief may be endorsement or explanation. The personal- or subpersonal-level processes of hypothesis generation and confirmation may be subject to one or another bias within the normal range, or to frank abnormalities.

We then confronted the fact that the second factor in the two-factor framework is poorly specified in terms of cognitive function. We noted that the case for rejecting a belief may sometimes depend primarily on the weight of disconfirming evidence and sometimes on the low prior probability of the belief being true. Thus, the answer to the question why the patient does not reject the belief may vary in its details and we suggested that this may be another locus of parametric variation. An *a priori* analysis of belief evaluation, drawing on the normative standards of probability theory, suggested that the task of evaluating an accepted hypothesis requires a step back from the initial acceptance, a step in the direction opposite to a prepotent response or tendency. We arrived at an initial suggestion that belief evaluation involves working memory and executive processes of inhibition.

Dual-process accounts of reasoning distinguish between heuristic (System 1) processes and analytic (System 2) processes and a central idea is that the two kinds of processes can come into conflict. We suggested that belief evaluation involves two kinds of analytic processes, assessment of competing hypotheses and inhibition of imperatives (the imperatives of observational adequacy, explanatory adequacy, and conservatism). Experimental and neuroimaging studies of the belief-bias effect were then interpreted as supporting the proposal that the second factor in the explanation of delusions is an impairment of working memory or executive function with a neural basis in damage to the right frontal region of the brain.

We began the chapter with examples of anosognosia considered as a delusion and we ended by addressing two problems – how can anosognosia fit the two-factor framework and how could recognition of paralysis depend on working memory or executive function? In response to the first problem, we argued that a first factor impairs the patient's concurrent awareness of paralysis and a second factor explains why the patient does not make appropriate use of a mass of other available evidence. We approached the second problem theoretically by drawing on the analogy between belief evaluation and the task used in belief-bias experiments. In the presence of a first factor, there is a conflict between the response based on available evidence and the response based on prior beliefs. Under these conditions, impairments of working memory and executive function may have the consequence that prior beliefs that are now false are maintained rather than rejected. Finally, we presented some results from a study of cognitive impairments in right-hemisphere stroke patients. These results support our proposal about the nature of the second factor.

We are still some way from having a satisfactory explanatory account of even one pathology of belief. Much more needs to be discovered about the specific ways in which working memory and executive function are impaired in patients with different delusions. There may be further parametric variation depending on which of the imperatives need to be resisted. Halligan and Marshall may be right that, given a demanding notion of unification, 'it is unlikely that a unified theory of delusions will be forthcoming' (1996, p. 8). But we maintain the hope that the two-factor framework can be developed so that, on the one hand, it is not hopelessly underspecified and, on the other hand, it reveals a common structure in our explanations of pathologies of belief.

Dedication

This chapter is dedicated to the memory of John C. Marshall (1939–2007) whose foundational contributions to cognitive neuropsychology and cognitive neuropsychiatry continue to guide research in those fields. We remember with gratitude, and also with sadness, our conversations with him on the topics of this chapter.

Acknowledgements

At many points in this chapter we have drawn on the work of Max Coltheart and Tony Stone and we acknowledge, with thanks, their influence on our thinking about delusions. Thanks also to Rebekah White, and to the editors, Lisa Bortolotti and Matthew Broome, for their comments on an earlier version.

References

Adolphs, R., Tranel, D., Bechara, A., Damasio, H., and Damasio, A. R (1996). Neuropsychological approaches to reasoning and decision-making. In *Neurobiology of Decision-Making* (eds. A. R. Damasio, H. Damasio, and Y. Christen), pp. 157–179. Berlin, Springer-Verlag.

Aimola, A. M. (1999). *Dark Side of the Moon: Studies in Unilateral Neglect*. PhD Dissertation, University of Auckland,.

Aimola Davies, A. M., Davies, M., Ogden, J. A., Smithson, M., and White, R. C. (2009). Cognitive and motivational factors in anosognosia. In *Delusions and Self-Deception: Affective Influences on Belief-Formation* (eds. T. Bayne, and J. Fernandez), pp. 187–225. Hove, East Sussex, Psychology Press.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*. Washington, DC, American Psychiatric Association.

Arthur, A. Z. (1964). Theories and explanations of delusions: a review. *American Journal of Psychiatry,* **121**, 105–115.

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

Bentall, R. P., Corcoran, R., Howard, R., Blackwood, N., and Kinderman, P. (2001). Persecutory delusions: a review and theoretical integration. *Clinical Psychology Review,* **21**, 1143–1192.

Berti, A., Bottini, G., Gandola, M., *et al.* (2005). Shared cortical anatomy for motor awareness and motor control. *Science,* **309**, 488–491.

Berti, A., Làdavas, E., and Della Corte, M. (1996). Anosognosia for hemiplegia, neglect dyslexia, and drawing neglect: clinical findings and theoretical considerations. *Journal of the International Neuropsychological Society,* **2**, 426–440.

Berti, A., Làdavas, E., Stracciari, A., Giannarelli, C., and Ossola, A. (1998). Anosognosia for motor impairment and dissociations with patients' evaluation of the disorder: theoretical considerations. *Cognitive Neuropsychiatry,* **3**, 21–44.

Bisiach, E. (1988) Language without thought. In *Thought Without Language* (ed. L. Weiskrantz), pp. 464–484. Oxford, Oxford University Press.

Bisiach, E. and Geminiani, G. (1991). Anosognosia related to hemiplegia and hemianopia. In *Awareness of Deficit after Brain Injury: Clinical and Theoretical Issues* (eds. G. P. Prigatano, and D. L. Schacter), pp. 17–39. Oxford, Oxford University Press.

Bisiach, E., Vallar, G., Perani, D., Papagno, C., and Berti, A. (1986). Unawareness of disease following lesions of the right hemisphere: anosognosia for hemiplegia and anosognosia for hemianopia. *Neuropsychologia,* **24**, 471–482.

Blackwood, N. J., Howard, R. J., Bentall, R. P., and Murray, R. M. (2001). Cognitive neuropsychiatric models of persecutory delusions. *American Journal of Psychiatry,* **158**, 527–539.

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

Bortolotti, L. and Broome, M. R. (in press). A role for ownership and authorship in the analysis of thought insertion. *Phenomenology and the Cognitive Sciences*.

Bottini, G., Bisiach, E., Sterzi, R., and Vallar, G. (2002). Feeling touches in someone else's hand. *NeuroReport,* **13**, 249–252.

Breen, N., Caine, D., and Coltheart, M. (2002). The role of affect and reasoning in a patient with a delusion of misidentification. *Cognitive Neuropsychiatry,* **7**, 113–138.

Broome, M. R. and McGuire, P. K. (2008). Imaging and delusions. In *Persecutory Delusions: Assessment, Theory, and Treatment* (eds. D. Freeman, R. P Bentall, and P. A. Garety), pp. 281–301. Oxford, Oxford University Press.

Broome, M. R., Woolley, J. B., Johns, L. C., *et al.* (2005a). Outreach and support in south London (OASIS): implementation of a clinical service for prodromal psychosis and the at risk mental state. *European Psychiatry,* **20**, 372–378.

Broome, M. R., Woolley, J. B., Tabraham, P., *et al.* (2005b). What causes the onset of psychosis? *Schizophrenia Research,* **79**, 23–34.

Broome, M. R., Johns, L. C., Valli, I., *et al.* (2007). Delusion formation and reasoning biases in those at clinical high risk for psychosis. *British Journal of Psychiatry,* **191**(suppl. 51), s38–s42.

Bruce, V. and Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology,* **77**, 305–327.

Capgras, J. and Reboul-Lachaux, J. (1923). L'illusion des 'sosies' dans un délire systématisé chronique. *Bulletin de la Socété Clinique de Médicine Mentale,* **2**, 6–16.

Carpenter, K., Berti, A., Oxbury, S., Molyneux, A. J., Bisiach, E., and Oxbury, J. M. (1995). Awareness of and memory for arm weakness during intracarotid sodium amytal testing. *Brain,* **118**, 243–251.

Chatterjee, A. and Mennemeier, M. (1996). Anosognosia for hemiplegia: patient retrospections. *Cognitive Neuropsychiatry,* **1**, 221–237.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York, Praeger.

Cinan, S. and Öktem Tanör, Ö. (2002). An attempt to discriminate different types of executive functions in the Wisconsin Card Sorting Test. *Memory,* **10**, 277–289.

Colbert, S. M. and Peters, E. R. (2002). Need for closure and jumping-to-conclusions in delusion-prone individuals. *Journal of Nervous and Mental Disease,* **190**, 27–31.

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In *Attention and Performance XI* (eds. M. I. Posner, and O. S. M. Marin), pp. 3–37. Hillsdale, NJ, Lawrence Erlbaum Associates.

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

Coltheart, M. (2006a). John Marshall and the cognitive neuropsychology of reading. *Cortex,* **42**, 855–860.

Coltheart, M. (2006b). Acquired dyslexias and the computational modelling of reading. *Cognitive Neuropsychology,* **23**, 3–12.

Coltheart, M. (2007). Cognitive neuropsychiatry and delusional belief. *Quarterly Journal of Experimental Psychology,* **60**, 1041–1062.

Cotard, J. (1882). Du délire des négations. *Archives de Neurologie,* **4**, 152–170, 282–295.

David, A. S. (1993). Cognitive neuropsychiatry. *Psychological Medicine,* **23**, 1–5.

David, A. S. and Halligan, P. W. (1996). Editorial. *Cognitive Neuropsychiatry,* **1**, 1–3.

Davies, M. (2009). Delusion and motivationally biased belief: self-deception in the two-factor framework. In *Delusions and Self-Deception: Affective Influences on Belief-Formation* (eds. T. Bayne, and J. Fernandez), pp. 71–86. Hove, East Sussex, Psychology Press.

Davies, M. and Coltheart, M. (2000). Introduction: pathologies of belief. *Mind and Language,* **15**, 1–46.

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

De Neys, N. (2006). Dual processing in reasoning: two systems but one reasoner. *Psychological Science,* **17**, 428–433.

Edelstyn, N. M. J. and Oyebode, F. (1999). A review of the phenomenology and cognitive neuropsychological origins of the Capgras syndrome. *International Journal of Geriatric Psychiatry,* **14**, 48–59.

Ellis, H. D. (1998). Cognitive neuropsychiatry and delusional misidentification syndromes: an exemplary vindication of the new discipline. *Cognitive Neuropsychiatry*, **3**, 81–90.

Ellis, H. D. and Lewis, M. B. (2001). Capgras delusion: a window on face recognition. *Trends in Cognitive Sciences,* **5**, 149–156.

Ellis, H. D. and Young, A. W. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry,* **157**, 239–248.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science,* **11**, 19–23.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General,* **128**, 309–331.

Evans, J. St. B. T. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences,* **7**, 454–459.

Evans, J. St. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking and Reasoning,* **13**, 321–339.

Evans, J. St. B. T., Barston, J. L., and Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition,* **11**, 295–306.

Feinberg, T. E., Roane, D. M., and Ali, J. (2000). Illusory limb movements in anosognosia for hemiplegia. *Journal of Neurology, Neurosurgery, and Psychiatry,* **68**, 511–513.

Fine, C., Craigie, J., and Gold, I. (2005). Damned if you do, damned if you don't: the impasse in cognitive accounts of the Capgras delusion. *Philosophy, Psychiatry, and Psychology,* **12**, 143–151.

Fine, C., Gardner, M., Craigie, J., and Gold, I. (2007) Hopping, skipping or jumping to conclusions? Clarifying the role of the JTC bias in delusions. *Cognitive Neuropsychiatry,* **12**, 46–77.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA, MIT Press.

Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA, MIT Press.

Fodor, J. A. (1989). Why should the mind be modular? In *Reflections on Chomsky* (ed. A. George), pp. 1–22. Oxford, Blackwell.

Fodor, J. A. (2000). *The Mind Doesn't Work That Way*. Cambridge, MA, MIT Press.

Freeman, D. (2007). Suspicious minds: the psychology of persecutory delusions. *Clinical Psychology Review,* **27**, 425–457.

Freeman, D., Garety, P. A., Fowler, D., Kuipers, E., Bebbington, P. E., and Dunn, G. (2004). Why do people with delusions fail to choose more realistic explanations for their experiences? An empirical investigation. *Journal of Consulting and Clinical Psychology,* **72**, 671–680.

Freeman, D., Garety, P. A., McGuire, P., and Kuipers, E. (2005). Developing a theoretical understanding of therapy techniques: an illustrative analogue study. *British Journal of Clinical Psychology,* **44**, 241–254.

Frith, C. D., Blakemore, S.- J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* **355**, 1771–1788.

Garety, P. A. and 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., and Wessely, S. (1991). Reasoning in deluded schizophrenic and paranoid patients: biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease,* **179**, 194–201.

Gialanella, B. and Mattioli, F. (1992). Anosognosia and extrapersonal neglect as predictors of functional recovery following right hemisphere stroke. *Neuropsychological Rehabilitation,* **2**, 169–178.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist,* **46**, 107–119.

Gilbert, D. T., Krull, D. S., and Malone, P. S. (1990). Believing the unbelievable: some problems in the rejection of false information. *Journal of Personality and Social Psychology,* **59**, 601–613.

Gilbert, D.T., Tafadori, R. W., and Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology,* **65**, 221–233.

Goel, V. and Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition,* **87**, B11–B22.

Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines,* **8**, 101–118.

Halligan, P. W. and Marshall, J. C. (eds.) (1996). *Method in Madness: Case Studies in Cognitive Neuropsychiatry*. Hove, East Sussex, Psychology Press.

Halligan, P. W., Marshall, J. C., and Wade, D. T. (1995). Unilateral somatoparaphrenia after right hemisphere stroke: a case description. *Cortex*, **31**, 173–182.

Harris, M. E. (1988). *Wisconsin Card Sorting Test Scoring Program (Version 2.0)*. Odessa, FL, Psychological Assessment Resources, Inc.

Hasson, U., Simmons, J. P., and Todorov, A. (2005). Believe it or not: on the possibility of suspending belief. *Psychological Science,* **16**, 566–571.

Heilman, K. M. (1991). Anosognosia: possible neuropsychological mechanisms. In *Awareness of Deficit after Brain Injury: Clinical and Theoretical Issues* (eds. G. P. Prigatano, and D. L. Schacter), pp. 53–62. New York, Oxford University Press.

Heilman, K. M., Barrett, A. M., and Adair, J. C. (1998). Possible mechanisms of anosognosia: a defect in self-awareness. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **353**, 1903–1909.

Hemsley, D. R. and Garety, P. A. (1986). The formation and maintenance of delusions: a Bayesian analysis. *British Journal of Psychiatry,* **149**, 51–56.

Hohwy, J. and Rosenberg, R. (2005). Unusual experiences, reality testing and delusions of alien control. *Mind and Language,* **20**, 141–162.

House, A. and Hodges, J. (1988). Persistent denial of handicap after infarction of the right basal ganglia: A case study. *Journal of Neurology, Neurosurgery, and Psychiatry,* **51**, 112–115.

Humphreys, G. W. (1991). Review of Shallice, *From Neuropsychology to Mental Structure. Mind and Language,* **6**, 202–214.

Huq, S. F., Garety, P. A., and Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *Quarterly Journal of Experimental Psychology,* **40A**, 801–812.

Jaspers, K. (1963). *General Psychopathology* (Trs J. Hoenig and M. W. Hamilton). Manchester, Manchester University Press.

Jeannerod, M. and Pacherie, E. (2004). Agency, simulation and self-identification. *Mind and Language,* **19**, 113–146.

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry,* **160**, 13–23.

Kapur, S. (2004). How antipsychotics become anti-'psychotic': from dopamine to salience to psychosis. *Trends in Pharmacological Sciences,* **25**, 402–406.

Kapur, S., Mizrahi, R., and Li, M. (2005). From dopamine to salience to psychosis. Linking biology, pharmacology and phenomenology of psychosis. *Schizophrenia Research,* **79**, 59–68.

Karnath, H-O., Baier, B., and Nägele, T. (2005). Awareness of the functioning of one's own limbs mediated by the insular cortex? *Journal of Neuroscience,* **25**, 7134–7138.

Klauer, K. C., Musch, J., and Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review,* **107**, 852–884.

Levine, D. N. (1990). Unawareness of visual and sensorimotor defects: a hypothesis. *Brain and Cognition,* **13**, 233–281.

Lezak, M. D., Howieson, D. B., and Loring, D. W. (2004). *Neuropsychological Assessment*, 4th edn. New York, Oxford University Press.

Lie, C-H., Specht, K., Marshall, J. C., and Fink G. R. (2006). Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test, *NeuroImage,* **30**, 1038–1049.

Linney, Y. M., Peters, E. R., and Ayton, P. (1998). Reasoning biases in delusion-prone individuals. *British Journal of Clinical Psychology,* **37**, 285–302.

Lipton, P. (2004). *Inference to the Best Explanation*, 2nd edn. London, Routledge.

Maguire, A. M. and Ogden, J. A. (2002). MRI brain scan analyses and neuropsychological profiles of nine patients with persisting unilateral neglect. *Neuropsychologia,* **40**, 879–887.

Maher, B. A. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology* **30**, 98–113.

Maher, B. A. (1988). Anomalous experience and delusional thinking: the logic of explanations. In *Delusional Beliefs* (eds. T. F. Oltmanns, and B. A. Maher), pp. 15–33. Chichester, Wiley and Sons.

Maher, B. A. (1992). Delusions: contemporary etiological hypotheses. *Psychiatric Annals,* **22**, 260–268.

Maher, B. A. (1999). Anomalous experience in everyday life: its significance for psychopathology. *The Monist,* **82**, 547–570.

Marcel, A. J., Tegnér, R., and Nimmo-Smith, I. (2004). Anosognosia for plegia: specificity, extension, partiality and disunity of bodily unawareness. *Cortex,* **40**, 19–40.

Marshall, J. C. and Newcombe, F. (1966). Syntactic and semantic errors in paralexia. *Neuropsychologia,* **4**, 169–176.

Marshall, J. C. and Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research,* **2**, 175–199.

McKay, R. and Cipolotti, L. (2007). Attributional style in a case of Cotard delusion. *Consciousness and Cognition,* **16**, 349–359.

Mele, A. R. (2009). Self-deception and delusions. In *Delusions and Self-Deception: Affective Influences on Belief-Formation* (eds. T. Bayne and J. Fernandez), pp. 55–70. Hove, East Sussex, Psychology Press.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., and Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General,* **130**, 621–640.

Moran, R. (2001). *Authority and Estrangement: An Essay on Self-knowledge*. Princeton, NJ, Princeton University Press.

Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology,* **2**, 175–220.

Pryor, J. (2004). What's wrong with Moore's argument? *Philosophical Issues*, **14**, 349–378.

Robertson, I. H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (1994). *The Test of Everyday Attention (TEA)*. Bury St Edmunds, Thames Valley Test Company.

Schuurmans-Stekhoven, J. and Smithson, M. (submitted). Orientation to uncertainty as a possible dual process.

Smith, E. E. and Kosslyn, S. M. (2007). *Cognitive Psychology: Mind and Brain*. Upper Saddle River, NJ, Pearson Prentice Hall.

Smithson, M., Aimola Davies, A. M., and Davies, M. (submitted). Additional methods for comparing cases with control samples and testing for dissociation.

Stanovich, K. E. (1999). *Who is Rational? Studies of Individual Differences in Reasoning*. Mahwah, NJ, Lawrence Erlbaum Associates.

Stanovich, K. E. (2003). The fundamental computational biases of human cognition: Heuristics that (sometimes) impair decision making and problem solving. In *The Psychology of Problem Solving* (eds. J. E. Davidson and R. J. Sternberg), pp. 291–342. New York, Cambridge University Press.

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

Van Dael, F., Versmissen, D., Janssen, I. Myin-Germeys, I., van Os, J., and Krabbendam, L. (2006). Data gathering: biased in psychosis? *Schizophrenia Bulletin,* **32**, 341–351.

Vuilleumier, P. (2004). Anosognosia: the neurology of beliefs and uncertainties. *Cortex,* **40**, 9–17.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology,* **12**, 129–140.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology,* **20**, 273–281.

Wechsler, D. (1987). *Wechsler Memory Scale–Revised*. San Antonio, TX: The Psychological Corporation.

Weinstein, E. A. and Kahn, R. L. (1950). The syndrome of anosognosia. *Archives of Neurology and Psychiatry,* **64**, 772–791.

Weinstein, E. A. and Kahn, R. L. (1951). Patterns of disorientation in organic disease of the brain. *AMA Archives of Neurology and Psychiatry,* **65**, 533–534.

Weinstein, E. A. and Kahn, R. L. (1953). Personality factors in denial of illness. *AMA Archives of Neurology and Psychiatry,* **69**, 355–367.

Weinstein, E. A. and Kahn, R. L. (1955). *Denial of Illness: Symbolic and Physiological Aspects*. Springfield, IL, Charles C. Thomas.

Weinstein, E. A., Kahn, R. L., Malitz, S. and Rozanski, J. (1954). Delusional reduplication of parts of the body. *Brain,* **77**, 45–60.

Young, A. W. (2000). Wondrous strange: the neuropsychology of abnormal beliefs. *Mind and Language,* **15**, 47–73.

Young, A. W. and Leafhead, K. M. (1996). Betwixt life and death: case studies of the Cotard delusion. In *Method in Madness: Case Studies in Cognitive Neuropsychiatry* (eds. P. W. Halligan, and J. C. Marshall), pp. 147–171. Hove, East Sussex, Psychology Press.