

Introduction: Pathologies of Belief

MARTIN DAVIES AND MAX COLTHEART

Some people seem to believe quite extraordinary things. For example, people say:¹

- My closest relatives have been replaced by impostors.
- I am dead.
- I am being followed around by people who are known to me but who are unrecognizable because they are in disguise.
- The person I see in the mirror is not really me.
- A person I knew who died is nevertheless in the hospital ward today.
- This arm [the speaker's left arm] is not mine it is yours; you have three arms.
- Someone else is able to control my actions.
- Someone else's thoughts are being inserted into my mind.

These apparent claims are so bizarrely false that we might have some doubt whether the utterances can really be expressions of beliefs. But alternative interpretations of the utterances are not easy to sustain (see below, section 1.1). These people seem to understand what they are saying, they seem to be sincere,

We are grateful to all the participants in the workshop at Macquarie University at which the papers in this volume were first presented. Early versions of material in this Introduction were presented in talks at the Australian National University, at the Novartis Foundation in London and at the Institute of Psychiatry also in London; we learned much from discussions on those occasions. Special thanks to Mark Greenberg, Robyn Langdon, Christopher Peacocke, Ian Ravenscroft, Tony Stone and Andy Young for comments and conversations.

Address for correspondence: Martin Davies, Corpus Christi College, Oxford OX1 4JF, UK; Max Coltheart, Macquarie Centre for Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia.

Email: martin.davies@philosophy.ox.ac.uk; max@currawong.bhs.mq.edu.au.

¹ These eight examples are, respectively: the Capgras delusion (Capgras and Reboul-Lachaux, 1923; Young, this volume); the Cotard delusion (Cotard, 1882; Berrios and Luque, 1995; Young, this volume); the Fregoli delusion (Courbon and Fail, 1927; de Pauw, Szulecka and Poltock, 1987; Ellis, Whitley and Luauté, 1994); the delusion of mirrored-self misidentification (Foley and Breslau, 1982; Breen et al., this volume); a delusion of reduplicative paramnesia (Benson, Gardner and Meadows, 1976; Breen et al., this volume); a delusion sometimes found in patients suffering from unilateral neglect (Bisiach, 1988); and the delusions of alien control and of thought insertion, which are characteristic of schizophrenia (Frith, 1992).

and they sometimes act in ways that make sense in the light of the beliefs that they seem to be expressing. So there is a strong case for taking their utterances at face value. But if we do accept that these people are expressing genuine beliefs then, in each case, we are bound to ask how such a bizarrely false hypothesis could come to be believed.

The difficulty in taking these utterances as expressions of beliefs can be seen to flow from the fairly widely accepted idea that the attribution to people of beliefs is governed by a constraint of rationality or reasonableness.² It is not easy to say just what this constraint comes to and it would clearly not be right to insist that a proper interpretation of a person's mental life must make him or her out to be perfectly logical and rational at every point. But we normally expect that a person's beliefs, desires and intentions will exhibit a certain kind of intelligibility that is distinctive of the folk psychological domain of experiencing, thinking, reasoning and planning.

Christopher Cherniak (1986) argues against accounts of belief that impose an 'ideal general rationality condition'. Such a condition, he says, is too strong because it ignores the fact that 'human beings are in the *finitary predicament* of having fixed limits on their cognitive capacities and the time available to them' (p. 8). But he also maintains that the rationality requirements on belief attribution should not be too weak (1986, p. 6):

A cognitive theory with no rationality restrictions is without predictive content; using it, we can have virtually no expectations regarding a believer's behavior. There is also a further metaphysical, as opposed to epistemological, point concerning rationality as part of what it is to be a *person*: the elements of a mind . . . must *fit together* or cohere.

If we cannot make any sense at all of how a certain person could reasonably have arrived at a particular belief on the basis of experience and inference then this counts, provisionally even if not decisively, against the attribution of that belief to that person. One way that we can apply this test of intelligibility is to imagine ourselves in the other person's situation, or even imagine being the other person, and then to consider what, from that person's point of view, would be a reasonable thing to believe. We *simulate* the other person in imagination and seek to understand him or her 'from the inside'. This simulation methodology has its limitations. It seems possible in principle that someone might rationally arrive at a belief on the basis of an experience that is so strange, or so terrifying even to imagine, that it is not possible for us to bring

² This idea is familiar from the work of Davidson, 1984, on radical interpretation and Dennett, 1987, on the intentional stance. But even someone who thought that the connection between beliefs and rationality was just contingent should allow that we have to address the question how such a strange hypothesis could come to be believed. For the distinction between rationality and reasonableness, see e.g. Church, 1987.

ourselves fully to identify with that person in imagination. But, whether or not simulation is a good way of applying the intelligibility test, the first question to ask about the apparent beliefs with which we began is whether we can understand them as at least *prima facie* reasonable responses to experienced situations.³

The requirement of an intelligible link between belief and experience is one aspect of an overarching constraint of rationality or reasonableness. But we also expect to find a person's beliefs fitting together with each other tolerably well, particularly in the person's deductive and probabilistic reasoning. This is a second aspect of the constraint.

Given our finitary predicament, it is important not to overstate this expectation of overall consistency.⁴ If a person believes some proposition A, and B is a proposition that follows logically from A, we do not always expect him to believe B as well. Indeed, a person might believe A, recognize that A entails B, and yet believe not-B rather than B. Suppose, for example, that a thinker antecedently believes A and has also reached the view that B is false on the basis of what seem to be quite strong considerations. He now sees that there is a logical relation between A and B, so that he cannot be right both in his belief that A is true and in his view that B is false. Something needs to give. But the task of deciding which belief should be changed might be a difficult and complex one, and he might lack the time and resources to carry it out immediately. He might take the strategic decision to tolerate the unresolved inconsistency until time and resources become available. We should also note that our thinker might find the prospect of changing either of the incompatible beliefs deeply uncomfortable and so, quite apart from issues of time and resources, might be motivated to tolerate a cognitive dissonance (Festinger, 1957).

Our thinker might live, for the time being or even permanently, with an acknowledged tension in his system of beliefs. And if this is the case when the tension arises from a deductive relation then it is all the more so for tensions arising from probabilistic relations. In short, an overarching constraint of

³ We might be able to understand these apparent beliefs as *prima facie* reasonable responses to experienced situations even though, because of their implausibility, they are not really reasonable all things considered. This leads on to the second question in the text.

⁴ Dennett, 1981, p. 44, says, 'The belief store must be – in the main – consistent', and goes on to cast doubt on the idea that a 'neurocryptographer' could insert into someone's (Tom's) brain the single belief, 'I have an older brother living in Cleveland', where this is not only false but also inconsistent with much that Tom already believes. According to Dennett, the outcome of the neurocryptographer's efforts will be either that the candidate belief will be rejected by Tom on the grounds of its incompatibility with his other beliefs or else that Tom will end up endorsing contradictions, such as 'I am an only child and have an older brother living in Cleveland'. In the latter case, we are to conclude that Tom's basic rationality is impaired so that 'in neither case has our neurocryptographer succeeded in wiring in a new belief' (*ibid.*). Dennett's vivid example is clearly relevant to the topic of delusional beliefs.

rationality or reasonableness must allow that there are many ways to explain or excuse departures from the ideal of overall consistency in a system of beliefs.

We have said that we should not pitch too high the standard of overall consistency that is required for the attribution of beliefs. But there is still a second question to be faced if we are to regard those bizarrely false apparent beliefs as genuine. In addition to the question whether the hypothesis in question can be understood as a *prima facie* reasonable response to the subject's experience, we have to ask why the hypothesis is adopted and maintained as a belief despite its utter implausibility and the uniform scepticism with which other people greet it. We can establish that this question may have an answer by pointing out that departures from overall consistency do occur and can be excused. But an answer still needs to be given.

1. Understanding Delusions

If the examples with which we began are regarded as genuine beliefs then they would seem to be delusions according to the definition offered by the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, 1994, p. 765)*:

Delusion A false belief based on incorrect inference about external reality that is firmly sustained despite what almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary.

The eight examples of apparent beliefs are all, we assume, false and they are all, at least to some extent, about the external world rather than being exhaustively about the subject's own experience. They are adopted and maintained despite the best efforts of relatives, friends and medical staff to assure the subject that they are not true and despite the fact that they are massively implausible in the light of other things that the subject knows. The aspect of the *DSM-IV* definition about which we might be cautious at this stage is the idea that a delusion must be 'based on incorrect inference'. It is plausible that many delusional beliefs are arrived at by a route that involves a flawed inference, either deductive or inductive, from true premises about experience to a false conclusion about external reality. But we should also allow that some delusional beliefs might be arrived at just by taking an illusory experience to be veridical (see below, section 4.3). That need not, strictly speaking, involve a step of inference on the part of the subject since the subject might make a judgement about the external world without going via any premises about his or her own experience.⁵

⁵ For a helpful survey of definitions of delusion, see Garety and Hemsley, 1994, ch. 1.

Theorists – whether philosophers, psychologists, psychopathologists, or psychiatrists – have taken differing views of the extent to which delusions can be understood, or rendered intelligible, in folk psychological terms. It is often remarked, for example, that Karl Jaspers (1963) held that the primary delusions of schizophrenic patients are not understandable, by which he seems to have meant that they are not folk psychologically intelligible and cannot be understood empathetically ‘from the inside’. One gloss on Jaspers’s view would be that the onset of a delusion does not have a rational explanation but just a causal explanation in terms of some organic disease process. For these purposes, a rational explanation of a subject’s belief is an explanation that reveals the belief to be the right thing for the subject to think, or at least a reasonable thing for the subject to think, given the subject’s experiences and other beliefs. But the issues here are quite complex because the notions of a belief that is folk psychologically intelligible, a belief that can be understood empathetically, and a belief that has a rational explanation do not coincide perfectly.

We have already noted, in effect, that a belief might have a rational explanation in terms of a subject’s experience although the strange or terrifying nature of that experience might stand in the way of our identifying properly with the subject in imagination. It may also be that a belief is folk psychologically intelligible even though it does not have an explanation that is cast wholly in terms of right thinking. Our everyday psychological understanding includes some understanding of excusable departures from right thinking and some understanding of the not-ideally-rational strategies that we employ in the face of the finitary predicament.

In addition, a belief might be folk psychologically intelligible even though it cannot be wholly understood ‘from the inside’ just by imagining being in the subject’s situation. We understand, for example, that a person’s dispositions to form beliefs on the basis of evidence might be perturbed in the direction of excessive caution or excessive boldness by factors such as mood or by the operation of a drug. If we imagine ourselves in the situation of a subject whose thinking is being perturbed in one way or another then we may sometimes be able to simulate the subject’s thinking just by bringing to bear our own knowledge ‘from the inside’ of what it is like to be in that kind of situation. But there may be perturbing influences that we understand quite well even though we have never experienced them ourselves, and this understanding may fall within the scope of folk psychology. In these cases of folk psychological intelligibility, simply imagining ourselves in the subject’s situation is not enough; we have to rely on a body of third-personal knowledge ‘from the outside’ about the way in which the particular perturbing influence works.⁶

⁶ These two paragraphs show (i) that a belief may have a rational explanation but not be understandable by simulation; (ii) that a belief may be folk psychologically intelligible but not have a purely rational explanation; and (iii) that a belief may be folk psychologically intelligible but not be understandable purely by simulation. It is plausible that simulation of

Given these distinctions, it is possible to develop Jaspers's view in more than one way. Each possible development is potentially problematic in the light of a corresponding way of taking the initial idea that the attribution of beliefs is governed by a constraint of rationality or reasonableness. Thus, for example, if Jaspers's view is that delusions are not folk psychologically intelligible then a problem arises when this is combined with the idea that the attribution of beliefs is governed by an overarching constraint of intelligibility. The problem in this case is that, given the constraint on attribution, if the so-called delusions really defy folk psychological understanding then patients should not, after all, be credited with beliefs with those bizarrely false contents. In the same way, a problem arises if we combine the view that delusions do not have purely rational explanations with the idea that attributions of belief are governed by an overarching constraint of rationality, or if we combine the view that delusions cannot be understood 'from the inside' with the idea that an overarching constraint on belief attribution can be cast in terms of simulation. In each case, the upshot would be that we should not attribute delusional beliefs.⁷

In fact, given the distinctions that we have drawn, we have considerable room for manoeuvre to avoid this problem; and, in any case, we are not bound to accept Jaspers's view. So the option of attributing bizarrely false beliefs to patients surely remains open. But we shall briefly consider the alternatives to such attributions.

1.1 Reinterpreting Patients' Utterances

If we were not to attribute bizarrely false beliefs then the utterances that are apparent expressions of delusions would have to be given some other interpretation.

One possibility is that those utterances should not be regarded as expressions of the subject's thoughts at all but should be regarded, from a psychological point of view, as nothing more than noise.⁸ Another possibility is that they should be treated as metaphorical, or in some other way non-literal, expressions of beliefs that are not bizarrely false. But the 'mere noise' and 'non-literal' options do not seem to offer satisfying accounts of examples such as those with which we began. On the one hand, it does not seem right to take the extreme course of giving up on psychological interpretation as the 'mere noise' option suggests. On the other hand, it is inadequate to offer interpretations of patients'

flawed thinking by someone susceptible to the same error illustrates the converse of (i) and that the thinking of savants may illustrate the converse of (ii). Perhaps it is not obvious that the converse of (iii) is a possible combination. In any case, it is clear enough that the three notions are distinct.

⁷ See Eilan, 2000, for a discussion of the apparent paradox implicit in Jaspers's approach.

⁸ This is one way of taking the suggestion by Berrios, 1991, that the utterances of deluded patients should be regarded as 'empty speech acts'.

utterances as innocuous remarks to the effect that it is, in certain respects, *as if* their closest relatives had been replaced by impostors or *as if* they were dead.

There is a third possible way of avoiding the attribution to patients of bizarrely false beliefs, an alternative to both the 'mere noise' and the 'non-literal' options. The patients' utterances might be interpreted as expressions of mental states with those false contents, but not as expressions of beliefs. They might be expressions of imaginative mental states or (if this is different) of states of the kind involved in supposing that something is the case or in entertaining a hypothesis.⁹ As it stands, this third possibility, like the 'non-literal' option, seems to involve underdescribing the strangeness of the patients' psychological state. There is nothing at all pathological in merely imagining that one is being followed around by people who are known to one but who are unrecognizable because they are in disguise nor in toying with the supposition that the person one sees in the mirror is not really oneself. But the third possibility might be elaborated. It might lead to the idea that so-called delusional states belong to a category that is somehow intermediate between imagination and belief or to the proposal that subjects suffering from delusions may confuse what is imagined with what is believed.

The proposal that some delusions have their origin in a 'misidentification of imagination' (Currie, this volume) is an important one. There is considerable plausibility in the idea that some schizophrenic patients fail properly to distinguish between imagining something to be the case and believing it to be the case. A subject who imagines something to be so and then mistakes what is imagined for a belief may end up in a state that is intermediate between imagination and belief. Perhaps some so-called delusions in schizophrenia are states of this intermediate kind rather than bizarrely false beliefs. So the 'misidentification of imagination' approach can provide an alternative interpretation of some utterances that are apparent expressions of delusions. But there are other cases in which it seems plausible that a subject who starts by imagining that P and misidentifies this as a belief that P comes, in the end, genuinely to believe that P (Currie, this volume). So the 'misidentification of imagination' approach still allows that some schizophrenic patients have delusional beliefs.

We have been considering the idea that utterances that are apparent expressions of delusions should be given some other interpretation. The 'mere noise' and 'non-literal' options are not adequate. A third option is inadequate as it stands but can be elaborated into an approach that does suggest alternative interpretations for some apparent expressions of delusions in schizophrenia. However, even this 'misidentification of imagination' approach allows that some schizophrenic patients genuinely have delusional beliefs and, of course,

⁹ Gendler, 2000, argues that supposing for the sake of argument is different from, and does not involve, the kind of imagining that figures in make-believe.

it does not suggest any alternative interpretation in cases that are unrelated to schizophrenia.

In most cases, then, it seems that patients should be interpreted as having the bizarrely false beliefs that their utterances seem to express. So we return to the two questions that have to be asked about each of the eight examples with which we began. Can the delusional hypothesis be understood as a *prima facie* reasonable response to the subject's experience and how does the hypothesis come to be adopted and maintained as a belief despite its implausibility?

1.2 Delusions as Responses to Unusual Experiences

Everyday psychological understanding does not rule out the possibility of a belief that has a causal explanation but not a purely rational explanation. We ordinarily allow that beliefs may be produced by wishful thinking, subliminal advertising or hypnosis, for example. But intuitive notions of folk psychological intelligibility do struggle somewhat with the idea of a belief that is brutally caused by disease or injury. We can stay close to the intuitive conception of beliefs as being adopted on the basis of experience and inference if the element of brute causation enters one step earlier and disease or injury causes abnormal experiences. We can also avoid a struggle with intuitive notions of folk psychological intelligibility if disease or injury perturbs reasoning processes in ways that are not too far removed from the familiar effects of mood, tiredness or alcohol, for example.

Other theorists working at about the same time as Jaspers allowed that delusions might have their origins in perceptual disorders or in faulty reasoning.¹⁰ Indeed, Jaspers himself allowed that there are delusion-like ideas that can be seen as intelligible responses to aspects of experience such as affect, mood or hallucination. More recently Brendan Maher (1974, 1988, 1992) has defended the proposal that delusions are false beliefs that arise as rational responses to unusual experiences. Maher's proposal certainly addresses our first question about the way in which the delusional hypothesis arises. But it has little to offer in response to our second question about how the delusional hypothesis is actually adopted and maintained as a belief despite its implausibility and despite the scepticism of others.¹¹

For an account of delusions that treats them as folk psychologically intelli-

¹⁰ See Garety and Hemsley, 1994, ch. 6.

¹¹ It is natural to say that in accepting a delusional hypothesis despite its implausibility patients depart from norms of rationality. But it is not straightforward to say just which norms are being infringed. Gold and Hohwy, this volume, follow Lewis, 1986, in distinguishing between procedural rationality and content rationality and then argue that the irrationality of delusional beliefs in schizophrenia cannot be understood as a failure of either of these aspects of rationality. For a different distinction, between norms that relate to logical consistency and norms that govern the use and management of evidence, see Bermúdez, forthcoming.

gible responses to unusual experiences and offers answers to both our questions we turn to Tony Stone and Andrew Young's paper, 'Delusions and brain injury: The philosophy and psychology of belief' (1997). According to Stone and Young, many delusions arise from anomalous experiences that result from brain damage. Since the way that brain damage leads to anomalies in experience is not itself a matter for folk psychology, Stone and Young's account of delusions is not purely folk psychological. But they do regard the link between the unusual experience and the delusional belief as being folk psychologically intelligible.

Stone and Young offer their account as applying specifically to delusions that result from brain damage. Also, they are concerned with delusions that have two properties: they are *monothematic* – specific to a particular topic – and *circumscribed* – relatively unelaborated. (We can note that, in principle, these two properties are independent of each other. A subject might have a bizarrely false belief about just one specific topic but might follow out many of the consequences of this belief: monothematic but not circumscribed. Or a subject might have bizarrely false beliefs about many unrelated topics but not follow out the consequences of any of them: circumscribed but not monothematic.) Stone and Young contrast circumscribed monothematic delusions resulting from brain damage with some kinds of schizophrenic delusions which 'can be florid and wide-ranging, with patients seeming to produce a new delusion in answer to almost every question, and effectively living in a solipsistic delusional world' (1997, p. 329). Some schizophrenic patients exhibit circumscribed monothematic delusions that may well fall within the scope of Stone and Young's account, but that account is not intended to apply to the extreme cases of florid schizophrenic delusions.¹²

2. The Capgras Delusion

The key idea in Stone and Young's account of monothematic and circumscribed delusions that result from brain injury is that (1997, p. 330):

these delusions can be best explained in terms of the person suffering from the delusion trying to make sense of or explain a disturbing perceptual experience that is brought about by the brain injury. On this view, the brain injury does not alter beliefs directly, but only indirectly by affecting the person's perceptual experiences.

¹² Stone and Young's (1997) account begins from the idea that patients have anomalous experiences as a result of brain damage. See Breen et al., this volume, for discussion of the comparison between delusions that occur as part of a primary psychiatric disorder, such as schizophrenia, and delusions that occur as the result of an identifiable neurological correlate.

The account is illustrated in terms of the explanation of the Capgras delusion that Young and his colleagues have developed.¹³

Patients who suffer from the Capgras delusion believe that someone close to them, often a close relative, usually their spouse in the first instance, has been replaced by an impostor who looks just like the replaced person. This is not an especially rare condition. The first major report of the delusion was in 1923 and a recent review surveys 174 published cases.¹⁴ The Capgras delusion is monothematic and it is also relatively circumscribed. If the person in front of you is not your spouse of many years then a remarkable switch must have taken place; but how was the trick turned? Your spouse must be somewhere else, but where? And is your spouse alive or dead? The case is of evident interest to the police; have they been informed? Capgras patients sometimes elaborate their delusion to the extent of invoking some piece of technology, perhaps robots or, in a biotechnological age, clones. Sometimes they express an attitude of antagonism or of friendliness towards the impostor and act in a way that is intelligible in the light of that attitude (sometimes with tragic results). But they are remarkably uninterested in the location or fate of their spouse and do not initiate a search or report the disappearance to the police. Capgras patients do not seem to incorporate the consequences of their belief into their general account of how the world works.

Until recently, the Capgras delusion was often explained in psychodynamic terms. On one hypothesis, the delusion serves to resolve conflicting feelings of love and hate towards a close relative; an impostor can be hated without guilt. Alternatively, Oedipal guilt can be alleviated if the attractive person who looks just like your mother is not in fact a relative of yours at all.¹⁵ But wholly psychodynamic accounts are cast into doubt by the fact that neuropsychological investigations of patients with the Capgras delusion have shown evidence of right hemisphere damage in many cases.¹⁶ It would be highly surprising that so many Capgras patients should show brain damage, and that a vast majority of those showing brain damage should have damage to the right hemisphere, if the explanation of their delusion were to be wholly psychodynamic.

The explanation of the Capgras delusion developed by Young and his colleagues postulates that the delusion arises from a deficit in face processing that is a kind of mirror image of prosopagnosia. Prosopagnosic patients are unable to recognize familiar faces but they still show autonomic affective responses, as indicated by increased skin conductance, to familiar faces. So it is proposed

¹³ See Ellis and Young, 1990; Ellis and de Pauw, 1994; Ellis et al., 1997; Young, this volume. As we shall see at the end of this section, Stone and Young hold that this key idea about the patient trying to make sense of an unusual experience needs to be augmented by an appeal to reasoning biases.

¹⁴ Capgras and Reboul-Lachaux, 1923; Förstl et al., 1991.

¹⁵ For psychodynamic accounts, see Capgras and Carrette, 1924; Enoch and Trethowan, 1991.

¹⁶ See Breen et al., this volume, for a review of findings about right hemisphere abnormalities.

that Capgras patients have an intact face recognition system but a loss of affective responses to familiar faces. As a result (Stone and Young, 1997, p. 337):

[They] are subject to an anomalous perceptual experience – an experience of seeing a face that looks just like their relative, but without experiencing the affective response that would normally be part and parcel of that experience.

The delusion itself arises, according to this proposal, from an attempt on the part of the patient to explain this peculiar experience.

Some recent empirical evidence provides support for the proposal that the Capgras delusion arises from an affective deficit in face processing. Skin conductance responses were measured in Capgras patients who were shown familiar and unfamiliar faces. As predicted by the proposal, the increased response to familiar faces that is found in normal subjects was absent or greatly reduced in these patients (Ellis et al., 1997; Hirstein and Ramachandran, 1997).

But this affective deficit is not sufficient, by itself, to generate the delusion. There are patients in whom brain damage has led to a failure to discriminate autonomically between familiar and unfamiliar faces, but who are despite this not delusional (Tranel, Damasio and Damasio, 1995). Presumably, these non-delusional patients experience just the same absence of affective response to the faces of loved ones that is experienced by patients suffering from the Capgras delusion. More generally, it seems clear that someone could suffer from the affective deficit in face processing, and have anomalous experiences when seeing the faces of loved ones, without adopting the implausible hypothesis that spouse and relatives have been replaced by impostors. Such a subject might say that his experience of the faces of loved ones is strange, flat, not quite right, and even that it is *as if* these were not really his loved ones. He might agree that one possible explanation of what is going on that would be adequate to the nature of this particular experience would be that the perceived faces are not, in fact, the faces of his loved ones. In an imaginative spirit, he might entertain and explore that hypothesis. But he could still stop short of actually adopting and maintaining the explanatory hypothesis as a belief. It seems, then, that a perceptual or affective deficit resulting from brain injury is a factor in the aetiology of the Capgras delusion but that this deficit is not sufficient, by itself, to account for the delusion.

We find this pattern in the case of other delusions as well. A patient with right hemisphere brain damage may suffer from paralysis of his left limbs, may neglect objects and events in the left side of his world, and may have the delusional belief that his left arm is not his own (the sixth of the eight examples at the beginning). It seems plausible that the patient's paralysis is a factor in the aetiology of this delusion. But many patients with left-sided paralysis resulting from right hemisphere damage are not delusional (and do not show unilateral neglect). So the paralysis is not sufficient, by itself, to account for

the delusion.¹⁷ A schizophrenic patient may suffer from the delusion of alien control, believing that someone else can control his actions (the seventh of the eight examples at the beginning). It seems that an experience of alien control is a factor in the aetiology of this delusion. But patients with depersonalization disorder describe their experiences as being *as if* an alien is controlling their actions, yet these patients are not delusional. So an experience in which it feels as if one is not controlling the movement of one's arm, for example, is not enough to produce the delusional belief that one's arm is being controlled by someone else. The anomalous experience of alien control is not sufficient, by itself, to account for the delusion of alien control.¹⁸

In these cases, as in the case of the Capgras delusion, there must be other factors at work. Stone and Young themselves make a proposal about the nature of these further factors (1997, p. 341):

We . . . think that the perceptual deficit account [of the Capgras delusion] needs to be augmented by a theory of the reasoning biases that lead to the delusional interpretation of the perceptual anomalies produced by the perceptual deficit.

3. Reasoning Biases and Attributional Biases

In a discussion of the formation and maintenance of delusional beliefs, Richard Bentall (1994; see also Bentall and Kinderman, 1998) begins with a contrast between accounts that follow Maher in saying that delusions are rational interpretations of anomalous experiences and accounts that focus on the role of cognitive biases. He acknowledges that some delusions arise from anomalous experiences, including experiences that result from perceptual or affective deficits, as proposed, for example, by Young and his colleagues. But Bentall maintains that, nevertheless, persecutory delusions often occur without any perceptual or other experiential anomaly. For such delusions, Maher's approach is not appropriate and an account in terms of cognitive biases is obligatory.¹⁹

At the end of the previous section we reached the view (following Stone and Young) that, even where perceptual anomalies do occur, Maher's kind of

¹⁷ See below, section 6.6, for further discussion.

¹⁸ See below, section 7.1, and Langdon and Coltheart, this volume, for further discussion.

¹⁹ Bentall and Kinderman, 1998, p. 121: '[Maher's] account consists of two logically unrelated elements: first Maher proposes the positive hypothesis that delusions are always a reaction to some kind of unusual perception and, second, he proposes the negative hypothesis that delusions are *never* the product of abnormal (nonperceptual) cognitive processes.' Bentall and Kinderman's assessment is that Maher's positive hypothesis is not generally correct, but does offer a correct account of some cases such as the Capgras delusion. Their main point (as in Bentall, 1994) is to show that Maher's negative hypothesis is not well supported.

approach is not adequate by itself. It is true that the hypothesis that loved ones have been replaced by impostors provides one possible explanation of the Capgras patient's experience of familiar faces, and to that extent the delusional hypothesis offers a rational interpretation of that experience. But this does not explain why Capgras patients adopt and maintain the hypothesis as a belief, since many subjects share the experience but not the delusion and the hypothesis is intrinsically implausible. So we could hope that the biases to which Bentall appeals in his account of persecutory delusions where there is no experiential anomaly might also be of some help in augmenting the perceptual-affective deficit account of the Capgras delusion.

It is no part of Bentall's position that deluded subjects suffer from a gross and pervasive deficit in logical reasoning. But, he does draw attention to a body of experimental work that indicates that deluded patients perform differently from normal subjects on probabilistic reasoning tasks.²⁰ The basic finding from this research is that deluded subjects seek less information than normal controls do before reaching a judgement. In short, deluded subjects show a tendency to jump to conclusions. This is clearly a suggestive finding. Someone who jumps to conclusions might move too readily from the thought that the hypothesis that loved ones have been replaced by impostors provides one possible explanation of his experience to the judgement that this hypothesis is indeed true.

On the other hand, it may seem that this bias in probabilistic reasoning cannot be enough, by itself, to explain delusional beliefs. The performance of deluded subjects is, on average, closer to the Bayesian norms than the performance of normal subjects, who tend to be overly cautious. So it could appear unlikely that just this difference from the reasoning of normal subjects would lead into such bizarre falsehood. But there is, amongst deluded subjects, a substantial subgroup (between one third and a half) who are prepared to reach a conclusion when presented with just one item of evidence.²¹ Patients who show this pattern of performance with completely neutral material may, as Bentall and Kinderman say, be 'vulnerable to accepting ideas that seem nonsensical to others' (1998, pp. 122–3).

A bias in probabilistic reasoning does not, however, explain the particular theme or content of the hypothesis that a deluded patient adopts. So Bentall's account of persecutory delusions appeals, in addition, to biases in the kinds of explanations that subjects give for their own behaviour and the behaviour of other people: attributional biases. Empirical results in this area indicate that, by comparison with normal subjects, patients with persecutory delusions tend to blame other people when something goes wrong and tend to take the credit

²⁰ Huq, Garety and Hemsley, 1988; Garety, Hemsley and Wessely, 1991. The work is reviewed in Garety and Hemsley, 1994, ch. 7.

²¹ In the experiments in question, the prior probability of the hypothesis is 50 per cent and the probability of the hypothesis given the one item of evidence is 85 per cent.

themselves when something goes right.²² Patients with persecutory delusions thus show an exaggerated ‘self-serving bias’ of a kind that has sometimes been regarded as a mechanism for maintaining self-esteem.²³ There is also some evidence that patients with persecutory delusions show a bias in the deployment of attentional and memory resources in favour of stimuli that are related to threats. These patients are unable to avoid attending to the meanings of threat-related words. In a similar way, negative trait words are particularly salient for depressed patients. It is of particular interest that patients with persecutory delusions show this same effect with respect to negative trait words (Kinderman, 1994), a result that supports the idea that ‘deluded patients have an implicit, but explicitly denied, negative self-concept’ (Bentall and Kinderman, 1998, p. 130; see also Lyon, Kaney and Bentall, 1994).

On the basis of a wide-ranging review, Bentall assembles the following account of persecutory delusions (1994, p. 353):

[P]aranoic patients show a tendency to ‘jump to conclusions’ in situations requiring probabilistic reasoning, attend selectively to threatening events, and exhibit a characteristic pattern of attributions, most marked when they make external attributions for negative events. . . . The biases observed indicate that persecutory delusions may have the function of protecting the individual against chronic feelings of low self-esteem.

Let us now turn back from Bentall’s account of persecutory delusions that occur without any experiential anomaly to Stone and Young’s account of the Capgras delusion. Stone and Young note that ‘outside the topic of the [monothematic and circumscribed] delusion itself, patients may appear quite rational’ (1997, p. 329). These patients do not have major problems with logical reasoning. But it is open to Stone and Young to augment the perceptual-affective deficit account of the Capgras delusion by appealing to probabilistic reasoning, attentional and attributional biases. Indeed, because Capgras patients often exhibit suspiciousness and persecutory delusions, Stone and Young might appeal to some of the very same biases as those that figure in Bentall’s account. The attributional bias (‘external attributions for negative events’) would help to explain why the hypothesis that loved ones have been replaced by impostors is treated as a particularly salient candidate explanation of the Capgras patient’s experience. The probabilistic reasoning bias, which is exhibited even on tasks in which there is no connection with threat or persecution, would help to

²² See e.g. Kaney and Bentall, 1989, 1992.

²³ We are not committing ourselves to this claim about self-esteem – or even to the idea that the unrefined concept of self-esteem is a useful one. For different kinds of ‘self-representation’, see Higgins, 1987; and for a review of discrepancies between these self-representations in deluded patients, see Bentall and Kinderman, 1998.

explain why that hypothesis, once considered, is actually adopted as a belief. Then, if Bentall is right about the function of persecutory delusions, the maintenance of the Capgras delusion might be explained in motivational terms.²⁴

So, on Stone and Young's account, augmented by appeal to various biases, the Capgras delusion depends on contributions from two sets of factors. In the first set are perceptual or experiential abnormalities resulting from brain damage. The second set is a heterogeneous collection of factors that lead to a misinterpretation of an anomalous experience (1997, p. 344):

The unusual perceptual experiences are considered to follow directly from brain malfunction, but their misinterpretation may reflect a further consequence of brain disease, an exacerbation of pre-morbid dispositions, or an indirect result of other changes.

This leaves it open that, in Capgras patients, the misinterpretation of anomalous experience might result from a cognitive deficit; that is, from impairment to, or even abolition of, a particular component or module within the normal cognitive processing machinery.

In contrast, Bentall is explicit that he is proposing cognitive biases rather than deficits (1994, p. 353):

It is worth repeating that these are cognitive biases rather than deficits. There is no evidence that deluded patients are incapable of reasoning; rather, they tend to weigh evidence relevant to their beliefs in a different way than normal individuals.

It may not be clear just what the bias-versus-deficit distinction is supposed to amount to here.²⁵ But we can note that in the quoted passage Bentall does not provide a compelling general argument to rule out the possibility that deluded patients have a reasoning deficit. For it is not right to suppose that the only cognitive deficit that could be proposed would be the total abolition of the reasoning system. If the reasoning system were to be made up of many components, then it might be that impairment to, or abolition of, one component would have the result that subjects would become less cautious and would jump to conclusions. An analogy may be useful. Surface dyslexic patients are more likely to produce a regular pronunciation of irregular or exception words (reading 'pint' to rhyme with 'mint') than are normal subjects. But it would not be right to insist on an explanation of this disorder in terms of a cognitive

²⁴ The empirical results on probabilistic reasoning by deluded patients do not indicate tenacious maintenance of hypotheses. On the contrary, some of the deluded subjects were ready to abandon a hitherto well confirmed hypothesis on the basis of just one item of disconfirmatory evidence (Garety, Hemsley and Wessely, 1991).

²⁵ See below, section 5.1, for further discussion of biases and deficits.

bias rather than a cognitive deficit on the grounds that ‘there is no evidence that surface dyslexic patients are incapable of reading’.

The point that we have now reached, following Stone and Young, is that the Capgras delusion arises from an affective deficit in face processing together with other factors including attributional biases, motivational factors, and abnormalities of probabilistic reasoning. But Stone and Young also provide another way of thinking about biases in the system of belief formation and maintenance and this is the topic of our next section.

4. Observational Adequacy and Conservatism

Beliefs, once formed, have a kind of inertia. As we revise our beliefs, we prefer changes that require less rather than more disruption in our system of beliefs as a whole. We tend to reject hypotheses that are inconsistent with many things that we already believe, especially if the hypotheses clash with propositions that play a pivotal organizing role in the web of our beliefs. Stone and Young call this a principle of *conservatism*.

The most conservative strategy would be to avoid all change in our system of beliefs. Suppose that I antecedently believe that there are no mice in the Coombs Building and then have a perceptual experience that would be well explained by the hypothesis that there is a mouse in my office.²⁶ Adopting the ultra-conservative strategy, I could decline to believe that there is a mouse in my office and instead could try to generate an alternative explanation of my experience in a way that was consistent with my antecedently held beliefs. It is not really clear that any alternative explanation could consistently preserve all of my antecedent beliefs, but for present purposes the more important point is that, in general, it is unwise to refuse to alter one’s beliefs about the world in the light of one’s experiences. So there is another principle governing belief revision which Stone and Young call a principle of *observational adequacy*.

The idea of this second principle is that beliefs should be revised so as to be consistent with the observed data. Sitting in my office, sober and in good light, I seem to see a mouse in the corner. In general, I accept the deliverances of perceptual experience as veridical and the principle of observational adequacy dictates that my system of beliefs should be updated in the light of the observed datum that there is a mouse in the corner of my office. I should revise my belief that there are no mice in the Coombs Building.

But suppose that, sitting in my office, I seem to see in the corner several little green men playing blackjack with a pink elephant dealing the cards. It might be that I antecedently have an explicitly articulated and compelling set of reasons for believing that there are no blackjack-dealing pink elephants in

²⁶ The Philosophy Program, Research School of Social Sciences, Australian National University is located in the Coombs Building.

the Coombs Building. Or it might be that, although I have never set out to construct an explicit justification, the proposition that my office is not populated by green men and pink elephants figures as an implicit background assumption in my assessments of the way in which evidence confirms or disconfirms hypotheses. Either way, if I accept my experience as veridical and try to incorporate the 'little green men and a pink elephant' hypothesis into my system of beliefs then very substantial disruption will result. Here, the principle of conservatism outweighs the principle of observational adequacy and dictates that I should not take my experience at face value. I should deny the apparent data.

Conservatism and observational adequacy tend to pull in opposite directions. There is no general formula for balancing the two principles against each other. We shall not always do well to reject the deliverances of perceptual experience just because they clash with our antecedent beliefs and expectations; we have to learn from our experience. But we also understand that experience can sometimes be misleading.

4.1 Two Ways of Understanding Observational Adequacy

Usually, our attention is on the world that we perceive rather than on our experience of it. I used to believe that there were no mice in the Coombs Building; now I have to take account of the fact that there is one right here in my office. This is what the principle of observational adequacy requires.

But when, in the light of the principle of conservatism, we reject the deliverances of perceptual experience as misleading, our experience itself may become an object of enquiry. I still believe that there are no pink elephants in the Coombs Building; but I want to understand why it perceptually seemed to me that there was a pink elephant in the corner of my office. Now the data to be accounted for, according to the principle of observational adequacy, concern the character of my experience. The principle of conservatism dictates that the facts about my experience should be explained with the minimum of disruption to my system of beliefs; and we can suppose that this will turn out to be a relatively simple matter.²⁷

Experience itself may become the object of enquiry because it has been classified as illusory rather than accepted as veridical. But this is not the only

²⁷ It is easier to update my system of beliefs to accommodate the proposition that it misleadingly seemed to me that there was a pink elephant dealing blackjack than to accommodate the proposition that there really was a pink elephant there. But we must allow for the possibility that accounting for an illusion might turn out to be difficult. It is possible that an experience might first be classified as illusory, that updating the system of beliefs to account for the illusion might turn out to be massively disruptive, and that, in the end, the experience might be accepted as veridical after all. In such a case, a conclusion about the external world would be reached on the basis of detailed consideration of the character of my experience. This is quite different from the way in which, usually, I unreflectively take perceptual experiences at face value without directing attention on the experiences themselves.

way in which our explanatory interest can come to be fixed on the nature of our experience rather than on events in the external world. In the case of some experiences, there is no question of a correct or incorrect presentation of how things are in the external world. Itches and tickles, for example, are not classified as either veridical or illusory, but they can certainly claim the attention of the person undergoing them. The occurrence of such an experience may itself constitute data. Its occurrence may be quite unexpected; it may conflict with many things that I believe. In that case, the principle of observational adequacy dictates that the web of beliefs should be updated to take into account the data concerning the occurrence and nature of this experience. The principle of conservatism, as usual, requires that the updating should be done in such a way as to minimize overall doxastic disruption.

We can thus distinguish two slightly different ways of understanding the principle of observational adequacy. On the first construal, the observational data to which belief revision should be adequate concern the external world (e.g. there is a mouse in the corner of my office) rather than my experiences. On the second construal, the data to which belief revision should be adequate are data about my experiences (e.g. I have an itch; it seems to me that there is a pink elephant dealing blackjack). It will be helpful to keep this distinction in mind as we consider the proposal that Stone and Young make about biases in the system of belief formation and maintenance.

4.2 Two Ways of Interpreting Stone and Young's Suggestion

Stone and Young suggest (1997, p. 349) that, in cases of delusion, the balance between the principle of observational adequacy and the principle of conservativeness is tipped too far in favour of observational adequacy. There are two ways of interpreting this suggestion corresponding to the two ways of understanding observational adequacy.

According to Stone and Young, Capgras patients undergo an unusual kind of experience (1997, p. 337): 'an experience of seeing a face that looks just like their relative, but without experiencing the affective response that would normally be part and parcel of that experience'. This focus on the nature of the patients' experience makes it seem natural to interpret Stone and Young's suggestion in line with the second of the two construals that we just distinguished. But on further reflection it is not so clear that this interpretation makes best sense of Stone and Young's account of the Capgras delusion. It is not clear that Capgras patients attach too much weight to data that concern the nature of their experience.

A Capgras patient undergoes an anomalous experience and the hypothesis that a relative has been replaced by an impostor provides one possible explanation of the occurrence of an experience of this kind ('seeing a face that looks just like their relative, but without experiencing the affective response'). With respect to the occurrence and nature of the experience, a revision of the subject's beliefs to incorporate that hypothesis would be observationally adequate.

On the other hand, adopting that hypothesis and then following out all of its consequences would require a major upheaval in the subject's system of beliefs. Adopting the hypothesis as a belief would be very far from conservative and the judgement of a normal subject would be that the hypothesis should be rejected. But, the Capgras patient gives too little weight to conservatism and so adopts and maintains the hypothesis. It does not seem right, however, to say that the Capgras patient goes wrong by attaching too much weight to data about the nature of his experience. That anomalous experience does demand explanation. The Capgras patient's mistake is to be too ready to adopt a particular explanation of his experience, an explanation involving the delusional hypothesis rather than a more conservative alternative.

We suggest, therefore, that it is worth considering an interpretation of Stone and Young's suggestion in line with the first construal of the principle of observational adequacy so that the relevant data concern, not the subject's experience, but the external world. On this second interpretation,²⁸ a Capgras patient is too ready to accept a putative datum about the external world, a datum that is furnished by, but does not directly concern, the patient's anomalous experience. This interpretation would have to start from the idea that, when the patient sees a close relative, it seems to the patient that there is someone present who is not the relative but looks just like the relative.²⁹ The patient accepts this experience as veridical and so takes himself to be in possession of a piece of observational data; namely, that this person who looks just like the relative (and claims to be the relative) is not really the relative. Observational adequacy then requires a revision of the subject's system of beliefs to incorporate the hypothesis that this person who looks just like the relative is really an impostor. On the other hand, that would amount to a quite drastic breach of the principle of conservatism and the judgement of a normal subject would be that the hypothesis should be rejected. According to this interpretation of Stone and Young's suggestion, the Capgras patient attaches too much weight to the putative observational datum and not enough weight to conservatism and so adopts and maintains the delusional hypothesis.

A critic might say that there is something artificial about our distinction between two ways of construing observational adequacy and so also about our distinction between two interpretations of Stone and Young's suggestion. It might be said that our first construal of observational adequacy can be assimilated to the second. According to the first construal, the observational data

²⁸ Note that the *second* interpretation of Stone and Young's suggestion goes with the *first* construal of observational adequacy.

²⁹ This idea goes naturally with the thought that, in normal subjects, the affective response really is 'part and parcel' of the experience of recognizing a particular close relative or loved one. If the affective response normally contributes to the experience of visual recognition then the absence of the affective response would make a difference to a subject's experience of the world. The subject might say that, so far as the visual experience goes, the person no longer seems to be that particular close relative.

concern the external world; according to the second construal the data concern experience. For the second construal, we considered experiences, such as itches, for which the question of correct or incorrect presentation of events in the external world does not arise, and we considered experiences that have already been classified as illusory. But, it might be said, we could also consider experiences that are accepted as veridical. Instead of talking about the worldly datum that there is a mouse in the corner of my office, we could talk about the experiential datum that it seems to me that there is a mouse in the corner of my office. When this experience is accepted as veridical, the system of beliefs is updated in the light of acceptance of a particular explanation of the experience; namely, that things are as they appear to be and there really is a mouse in the corner of my office. Thus, it might be said, what we aim at, quite generally, is to update our system of beliefs so as to afford explanations of the occurrence and nature of our experiences (observational adequacy) and to do so in a way that minimizes doxastic disruption (conservatism). We do not need to separate out the first construal of observational adequacy on which the observational data concern the external world.

Similarly, our critic might continue, the second interpretation of Stone and Young's suggestion can be assimilated to the first.³⁰ On the second interpretation, we say that the Capgras patient gives too much weight to the putative worldly datum that this is not a relative but an impostor. But we can say, instead, that the Capgras patient gives too much weight to the experiential datum that it seems to him that this is not a relative but an impostor. However we have, in effect, already rejected this assimilation. It does not give an adequate account of the Capgras patient's error. There is, after all, nothing wrong with giving weight to that experiential datum. The patient's anomalous experience does demand explanation. But the correct explanation is that, as a result of brain injury, the patient is suffering from an affective deficit.

The critic who thinks that our distinctions are artificial may respond by saying that Stone and Young should be interpreted as suggesting that delusions arise from a bias in favour of accepting experiences as veridical. But now we no longer have any substantive disagreement with the critic. A bias in favour of accepting experiences as veridical can be described as going 'too far towards observational adequacy and against conservatism' (Stone and Young, 1997, p. 349). But this is only because the principle of observational adequacy dictates that belief revision should take account of the putative worldly data that a subject arrives at by accepting experiences as veridical. Our aim has been to highlight this construal of observational adequacy and to contrast it with the thinner idea that all that is required is that our system of beliefs should afford some explanation of the occurrence and nature of our experiences.

³⁰ Recall that the *second* interpretation of Stone and Young's suggestion goes with the *first* construal of observational adequacy.

4.3 Prioritizing the Delusional Hypothesis

Our reason for dwelling on the notion of observational adequacy is that the different interpretations of Stone and Young's suggestion involve different accounts of how the delusional hypothesis comes to be prioritized in the thinking of the Capgras patient.

When Stone and Young suggest that delusions arise from a bias in favour of observational adequacy and against conservatism it is natural to suppose that, on their view, the hypothesis that loved ones have been replaced by impostors scores higher than alternative hypotheses on observational adequacy. But, as we have seen, that hypothesis is no more observationally adequate to the nature of the Capgras patient's experience ('seeing a face that looks just like their relative, but without experiencing the affective response') than any of a host of alternative hypotheses. The observationally adequate alternatives include the correct hypothesis that the patient's experience is anomalous because of a brain injury. So, if we adopt the first interpretation of Stone and Young's suggestion (in line with the second construal of observational adequacy), then tipping the balance in favour of observational adequacy does not favour the delusional hypothesis over all others. It just leaves the delusional hypothesis as one candidate amongst many.

If we interpret Stone and Young's suggestion in this first way then the overall account of the formation of delusional beliefs needs to appeal to something like attributional biases to prioritize the delusional hypothesis over the other equally observationally adequate candidates (as suggested in section 3). The delusion could not be explained in terms of just two factors, a perceptual-affective deficit plus a bias towards observational adequacy at the expense of conservatism.

The situation is rather different if we adopt the second interpretation of Stone and Young's suggestion so that the observational data concern, not the subject's experience, but the external world. Recall that this interpretation starts from the assumption that the experience of a Capgras patient looking at the face of a close relative is not just anomalous but has a particular representational character. It seems to the patient that this person who looks just like the relative is not really the relative. It is then immediate that observational adequacy requires the adoption of beliefs that are consistent with the proposition that the person who is seen is not the relative despite looking just like (and claiming to be) the relative. So there is no evident need to appeal to attributional biases in order to prioritize the hypothesis that this person who looks just like the relative is really an impostor. The delusional hypothesis is prioritized by the default readiness to accept perceptual experiences as veridical and, to that extent, it is not ruled out that the delusion might be explained in terms of just the two factors that we mentioned.

If, however, the representational content of the patient's experience is less definite than we have assumed then accepting the experience as veridical will not, by itself, prioritize the delusional hypothesis. If, for example, it seems to

the patient merely that this is a person who looks just like the relative but leaves him feeling emotionally flat then some other factor will have to enter the account of how the specific delusional hypothesis comes to be prioritized.

We should also allow that, even when the delusional hypothesis is prioritized by the default readiness to accept perceptual experiences as veridical, attributional biases may enter the picture as factors in determining the representational character of the patient's experience itself. It may be only in the presence of some attributional bias in addition to the affective deficit that it would determinately seem to the patient that this person who looks just like the relative is *not* the relative.

5. Outline of an Account of Delusions

Stone and Young suggest that deluded patients show an abnormality in belief revision that can be described as a bias in favour of observational adequacy at the expense of conservatism. In this section, we consider more closely the nature and scope of the suggested abnormality in belief revision and then turn to the circumscription of delusions. We end the section with a summary of the various factors that might figure in the aetiology of a monothematic and circumscribed delusion following brain injury.

5.1 Belief Revision: Biases and Deficits

The most fundamental question about the abnormality in belief revision that is exhibited by deluded patients is, of course, what exactly it consists in. A secondary question that can help to organize the discussion is whether the abnormality is best conceived as a bias or as a deficit. We mentioned the bias-versus-deficit issue near the end of section 3. Bentall argues that the reasoning or belief revision abnormality in patients with persecutory delusions is a bias rather than a deficit on the grounds that the patients are not totally deficient in reasoning. But we noted that this does not, by itself, rule out the possibility that deluded patients have a reasoning deficit because impairment to, or abolition of, just one component of a complex cognitive system would normally be regarded as a cognitive deficit. We can clarify the issues here by distinguishing between two notions of deficit.

The notion of a deficit as an impairment to, or as the abolition of, a component of information-processing machinery is a familiar one, particularly in discussions of the cognitive psychological consequences of brain injury. But we can also allow that there is a notion of a deficit which, like the notion of a bias, figures in descriptions of particular patterns of behaviour. If a subject did not engage in any probabilistic reasoning at all then we might classify this as a total probabilistic reasoning deficit. If a subject's probabilistic reasoning was merely skewed in some way then this would be a probabilistic reasoning bias. Given these behavioural notions of both bias and deficit, it would be correct

to say (with Bentall) that deluded patients show a reasoning bias rather than a total reasoning deficit. But it would also be correct to point out that a bias in behaviour can sometimes be explained in terms of a deficit in the cognitive machinery underpinning that behaviour. (Surface dyslexia is evidenced as a bias towards over-regularization in the pronunciation of exception words. But that is quite consistent with its explanation being a deficit in the reading system consisting in impairment to, or abolition of, one component of that system.)

In investigations of patients following brain injury, a cognitive deficit is usually taken to be an information-processing impairment that is the result of brain damage. In that same context, the term 'bias' usually indicates something that was premorbidly present and is not related to the injury. So if, following brain injury, a patient shows a behavioural bias in reasoning, then this might be explained in terms of a cognitive deficit resulting from the injury or in terms of a condition that was already present before the injury.

According to Young and his colleagues, one factor in the Capgras delusion is an anomalous experience that results from brain injury but there are other factors at work as well. Bentall offers an account of persecutory delusions in patients without any brain injury and appeals to various reasoning and attributional biases. So the simplest extrapolation from Bentall's account of persecutory delusions would say that the additional factors at work in producing the Capgras delusion are premorbidly existing biases that are unrelated to the brain injury. What Stone and Young allow, however, is that the additional factors may include premorbid dispositions but may also include further consequences of brain damage (1997, p. 344). They allow, that is, that the explanation of the Capgras delusion may involve a second cognitive deficit.

We have seen that deluded patients do seem to show what might be called behavioural biases in their probabilistic reasoning. They tend to jump to conclusions. But what concerns us here is not just the initial adoption of the delusional hypothesis but also the robustness of the maintenance of that hypothesis in the face of conflicting considerations. It may be that for a period of weeks, months or even years, no amount of disconfirmatory evidence will persuade the patient to abandon the delusional belief. This is not, of course, a total deficit in reasoning or in belief revision. But nor is it a mere tendency to hold onto the hypothesis, or to require more disconfirmatory evidence than normal subjects do. Abandoning the delusional hypothesis is something that many patients simply do not do and so we are inclined to say that the abnormality in belief revision can be described in behavioural terms as a deficit, rather than just as a bias.³¹ It is, of course, a further question whether this

³¹ The notions of bias and deficit that are used in the description of patterns of behaviour are to some extent interchangeable. A pattern that is described as a bias, rather than as a total deficit in some domain such as reasoning or reading, may nevertheless be redescribed as a more specific deficit – as the absence of some particular aspect of normal reasoning or reading, for example.

We note here that the Capgras delusion is not always maintained over an extended period.

particular deficit in belief revision behaviour is to be explained in terms of an underlying deficit in some component of belief revision machinery.

Before we can pursue that further question, we need a fuller account of what it is that the deluded patient does not do. As we have already noted, it would not be correct to say that patients fail to manifest any understanding of inferential liaisons. Understanding of those liaisons is surely manifested in patients' anticipation that other people will find their beliefs massively implausible. Nor would it be correct to say that patients do not actually engage in any inferences involving their delusional hypothesis. Even where delusions are relatively circumscribed rather than florid, patients engage in some degree of elaboration. What patients do not do is to allow antecedent beliefs or stored knowledge to undermine their commitment to the delusional hypothesis.

It would not be right to describe this by saying that patients attach no weight at all to their antecedent beliefs or stored knowledge. Their life would be impossible unless they were sometimes prepared to reject a hypothesis on the grounds of conflict with something that they already knew. We might describe the situation by saying that patients attach no weight to their antecedent beliefs or stored knowledge *by comparison with* their delusional hypothesis. Equally, and more straightforwardly, we can say that they attach to their delusional hypothesis a weight greater than any other. But now we must ask what property of their delusional hypothesis it is that results in its being given this greatest weight.

5.2 Belief Revision: The Nature and Scope of the Abnormality

We have suggested that patients attach greatest weight to hypotheses in some class that includes the delusional hypothesis. In order to pursue the question what the defining characteristic of this class of hypotheses is, it will be helpful to recall the two ways of interpreting Stone and Young's suggestion about a bias in favour of observational adequacy at the expense of conservatism (section 4.2). Each interpretation of their suggestion leads to a possible answer to our question.

We said that on the first way of taking Stone and Young's suggestion, the overall account would need to appeal to something like attributional biases to prioritize the delusional hypothesis. So it might be that the hypotheses to which greatest weight is attached are those that are favoured by the attributional bias that prioritizes the delusional hypothesis. If this were the right answer to our question then it is not obvious that there would be any good

Mackie, Ebmeier and O'Carroll, 1994, p. 212, describe a case as follows: '[The patient] stated that over the past year people had "changed", particularly his parents and close relatives. They would often leave a room and an almost identical impostor would return. He could identify impostors, because they were marginally shorter and wider than real individuals. Such changes would occur repeatedly, often many times within the same day. The process angered him considerably, and he felt it was carried out in an attempt to "get" him.'

reason to postulate an underlying deficit in belief revision machinery in order to account for the maintenance of the delusional hypothesis in the face of conflicting considerations. As Bentall (1994) notes, there is some support for a link between attributional biases and motivation via the idea that attributional biases have the function of protecting self-esteem. So it would be natural to suppose that the maintenance of hypotheses that are favoured by an attributional bias would have a motivational explanation.

If anything like this were right then the overall account of the Capgras delusion would have two components. There would be the perceptual-affective deficit in face processing and there would be the motivated maintenance of a hypothesis favoured by an attributional bias. The apparent behavioural deficit in belief revision would simply be a consequence of this second attributional-motivational component rather than a manifestation of impairment to, or abolition of, some component of the cognitive machinery that underpins belief revision.

On the second way of interpreting Stone and Young's suggestion, making use of the idea of a default readiness to regard perceptual experiences as veridical, we can propose that the hypotheses to which greatest weight is attached are those that are the immediate deliverances of perceptual experience. If this were correct then we might reasonably postulate an underlying cognitive deficit in some component that is implicated in bringing antecedent beliefs or stored knowledge to bear on perceptually generated hypotheses (Langdon and Coltheart, this volume).

But this proposed answer to our question yields some very strong empirical predictions. A patient who is genuinely unable to allow antecedent beliefs or stored knowledge to trump the deliverances of perceptual experience will be irremediably taken in by every visual illusion.³² Suppose, for example, that a Capgras patient views an Ames room in which an adult is positioned at the taller corner and a child at the shorter corner. The proposed answer to our question seems to yield the prediction that the patient will maintain, in the face of all conflicting considerations, that the child is really taller than the adult.

This is just a hypothetical case and we do not have data to report. But we are very doubtful that a Capgras patient would inevitably be taken in by every illusion. So there is a serious problem for this second answer to our question. In order to account for the Capgras delusion, it must be that greatest weight is attached to the hypothesis that this person who looks just like the relative is not really the relative. But greatest weight must not be attached to the hypothesis that the child in the Ames room is taller than the adult. Yet both hypotheses result from accepting illusory experiences as veridical.

In response to this problem it may be said that there is an important differ-

³² Similar worries could be raised against the proposal that it is always the hypothesis that first comes to mind that has greatest weight attached to it.

ence between an experience of the Ames room illusion and the Capgras patient's anomalous experience of a close relative. When a subject views an Ames room, he or she has an illusory experience for a short time. But the Capgras patient's experience of a close relative remains anomalous over days, months or even years. If we say that the hypotheses to which greatest weight is attached are those that are the immediate deliverances of consistently repeated perceptual experiences then this seems to promise an account of the Capgras delusion without yielding any prediction about the patient inevitably being taken in by brief exposure to the Ames room illusion.

However, this revised version of the proposed answer to our question still yields quite strong empirical predictions. It is possible for a deluded patient to have repeated exposure to an illusion. Perhaps a Capgras patient has a picture of the Müller-Lyer illusion in a prominent position on his or her bedroom wall. In such a case, the prediction is that the patient will maintain that one line is longer than the other in the face of all conflicting considerations including assurances from friends and relatives that this is just an illusion and detailed accounts of the illusion from psychology textbooks. Once again we have no data to report. But we suggest that this prediction raises a doubt about even the revised version of the second answer to our question.

We have been trying to answer the fundamental question what the deluded patients' belief revision abnormality consists in. It is intuitively correct that deluded patients show a particular behavioural deficit in their belief revision and it is tempting to suppose that this is to be explained in terms of a deficit in some component of the cognitive machinery underpinning belief revision. But we have seen that it is not at all easy to describe the abnormality in a way that meets two criteria.

The first criterion is, of course, that the description of the abnormality should collect together just the cases in which a hypothesis is maintained in the face of all conflicting considerations. The suggestion that the abnormality consists in always attaching greatest weight to the immediate deliverances of (consistently repeated) perceptual experiences seems likely to fail on this criterion. The second criterion is that it should be plausible that the behavioural deficit is underpinned by a deficit in the cognitive machinery that is distinctively implicated in belief revision. The suggestion that the abnormality consists in attaching greatest weight to hypotheses that are favoured by a certain attributional bias seems to fail on this criterion. Unless the fundamental question can be answered in a way that meets the two criteria, doubt is cast on the idea that the particular deficit in belief revision behaviour is to be explained in terms of an underlying deficit in some component of belief revision machinery.

On the other hand, the association between delusions and damage to the right cerebral hemisphere offers some support to the conjecture that the failure to reject delusional hypotheses depends on a deficit in some right hemisphere cognitive function. If that conjecture were correct then the aetiology of many delusions would involve two cognitive deficits: one that causes an anomaly in

perceptual or affective experience and another that causes an abnormality in belief revision (see Langdon and Coltheart, this volume). What we have just seen, however, is that difficult questions remain about the prospects for such a two-deficit account.³³

5.3 Conservatism and Circumscription

Stone and Young suggest that deluded patients show a bias in favour of observational adequacy at the expense of conservatism. We have considered the notion of observational adequacy in some detail; now we turn to conservatism. Stone and Young take the principle of conservatism from Fodor, who says (1987, p. 63):

[B]elief fixation appears to be a *conservative* process: the goal of the game is to accomplish the maximum in accommodating data at the minimum cost in overall disturbance to previous cognitive commitments.

Their suggestion is that deluded patients adopt hypotheses as beliefs even though doing so goes against the principle of conservatism. So it might be thought that, when a delusional hypothesis is adopted as a belief, there is substantial disturbance to the patient's previous cognitive commitments. But, at least in the case of circumscribed delusions, this kind of doxastic disruption is not what happens. A Capgras patient, for example, typically does not follow through and adopt all or even many of the consequences of the delusional belief that loved ones have been replaced by impostors. To a considerable extent, the patient's previous cognitive commitments remain in place.

We need to clarify how it can be that a deluded patient gives insufficient weight to the principle of conservatism and yet does not suffer from substantial doxastic disruption. Ideally, when we adopt a new belief we also update our web of beliefs to maintain overall consistency. There will be more than one way of doing this, and the principle of conservatism says that updating should involve minimum overall disturbance. Sometimes, however, it may turn out that there is no way to adopt a new belief and maintain overall consistency without substantial disturbance to previous cognitive commitments. If the cost

³³ In this discussion we are supposing that, in normal subjects, there are domain-general processes of belief revision. We have not considered the alternative view according to which the adoption, maintenance and revision of beliefs is a function of domain-specific heuristics. On this view, overall consistency would normally be maintained to the extent that the various domain-specific heuristics do a good job of tracking the truth. There would be no holistic processes for evaluating hypotheses or for updating the web of beliefs. If this alternative view were correct, then the adoption and maintenance of a delusional hypothesis might simply result from the operation of a domain-specific heuristic on the input provided by an anomalous experience.

in terms of disturbance is too high, then this counts against adoption of the belief. This is how conservatism and observational adequacy can come into conflict.

When Stone and Young say that a delusional hypothesis scores low on conservatism they mean that adopting the hypothesis as a belief, and then making adjustments elsewhere in the light of logical and probabilistic relations, would involve considerable disturbance to the subject's antecedent system of beliefs. A normal subject who strikes the right balance between observational adequacy and conservatism would abandon the observationally adequate hypothesis rather than depart from conservatism to the extent of making all those adjustments. But it does not follow that a subject who adopts a delusional hypothesis will actually disturb his antecedent system of beliefs by making all the adjustments that would be required for overall consistency. For we noted at the outset that even a normal subject may very well live with an acknowledged tension in his system of beliefs. A normal subject might believe A and believe not-B while recognizing that A entails B. So too, a patient might adopt a delusional hypothesis but fail to make the revisions in his system of belief that would be needed to maintain overall consistency.

The principle of conservatism says that we should minimize doxastic disruption as we update the web of beliefs to maintain overall consistency; it brings together the ideals of minimal disruption and overall consistency. We give too little weight to conservatism if we adopt a belief that requires substantial disruption if consistency is to be maintained. But by moving away from the ideal of overall consistency, a deluded patient can avoid substantial doxastic disruption even while adopting a hypothesis that scores low on conservatism.³⁴

In the case of normal subjects, we said that a departure from the ideal of overall consistency might be the result of a strategic decision in the light of a shortage of the resources needed for a considered revision. Alternatively, it might be motivated by the prospect of the discomfort that would accompany either of the revisions that would improve consistency. In the case of deluded patients, the latter, motivational, style of explanation may seem more plausible than the former, strategic, one. For a Capgras patient, the belief revisions that would be required to maintain overall consistency given the hypothesis that loved ones have been replaced by impostors would surely be disruptive and uncomfortable. In the case of some other delusions, aiming for overall consistency and embracing the resulting doxastic disruption might even lead to madness.

To suggest a motivational explanation for the circumscription, rather than elaboration, of a delusion is to commit ourselves to the idea that the deluded subject has some understanding of the consequences that follow logically from

³⁴ In effect, there are three constraints on belief revision: observational adequacy, overall consistency, and minimal doxastic disruption. A circumscribed delusion involves preserving observational adequacy and minimal disruption at the expense of overall consistency.

his delusional belief. This is suggested, in any case, by the fact that deluded patients recognize that their beliefs will be found implausible by other people. So we assume that deluded patients appreciate that their beliefs are inconsistent with many deeply entrenched beliefs of other people and, indeed, with many beliefs to which they themselves are strongly committed as well. We might put this by saying that the circumscription of a delusion is a matter of a motivated limitation of inferential *performance* rather than of a lack of knowledge or *competence* concerning inferential relations.³⁵

5.4 Multiple Factors in the Aetiology of Delusions

We can now summarize the various factors that might figure in the aetiology of a monothematic and circumscribed delusion following brain injury. First, there is a cognitive deficit resulting from the brain injury and this gives rise to a perceptual, affective or other experiential anomaly.

Second, a hypothesis about the cause of this anomaly is generated. Some factor privileges a false but observationally adequate hypothesis that would normally be rejected on the grounds that it does not measure up to the principle of conservatism. This factor might be an attributional bias. Alternatively, it might simply be a default readiness to accept perceptual experiences as veridical. In addition, attentional factors or even psychodynamic factors could enter the story at this point.

Third, the hypothesis is adopted and maintained as a belief despite the fact that it is inconsistent with many other things that the subject believes including propositions that play a pivotal organizing role in the subject's web of beliefs. This total resistance against conflicting considerations is appropriately described, at a behavioural level, as a deficit in belief revision though it is not easy to characterize the deficit precisely. It is provisionally suggested that this deficit is to be explained in terms of impairment to, or even abolition of, some component of the cognitive system responsible for belief revision and that this damage is also the result of brain injury. However, there are some serious open questions about this suggestion. Alternative accounts would appeal again to attributional and motivational factors.

Fourth, adoption and maintenance of the delusional hypothesis is achieved without a major upheaval in the subject's belief system. This is not a result of total failure to appreciate the inferential liaisons of the delusional belief; nor is there a total failure to deploy the delusional hypothesis in reasoning. But elaboration of the delusion does not extend to the point where substantial

³⁵ On some 'functional role' accounts of the necessary conditions for possessing particular concepts or having thoughts with particular contents, what is required is (tacit) knowledge of (some of) the inferential liaisons in which thought contents stand. We intend our remarks about delusions to be consistent with such accounts.

doxastic disruption would result. It is suggested that this circumscription of the delusion is supported by motivational factors.³⁶

In even briefer form, the four steps in this schematic account are an anomalous Experience, a prioritized Hypothesis, the adoption of this hypothesis as a Belief, and then finally the Circumscription of the delusion within the subject's web of beliefs. We turn now to the application of this EHBC account.

6. Applying the Account

Throughout our discussion up to this point we have followed Stone and Young (1997) in using the Capgras delusion as our central example. In this section, we briefly review that case and then attempt to apply the four-step (EHBC) schematic account of the aetiology of delusions to five more examples from the list of eight with which we began.

6.1 The Capgras Delusion

A Capgras patient usually believes that his spouse or another close relative has been replaced by an impostor. In terms of the four steps in our schematic account, we can propose, first (E), that as a result of brain injury the patient has a disorder of face processing that produces anomalous experiences when he sees a familiar face. Second (H), a hypothesis is generated: This person that the patient is looking at is not his spouse despite looking very much like his spouse and claiming to be his spouse. This person is claiming to be someone that she is not and so is an impostor. This might be an explanatory hypothesis prioritized by an attributional bias or it might simply be the result of the patient accepting an anomalous experience as veridical. Third (B), this hypothesis is adopted and maintained as a belief as the result of a deficit in belief revision that Stone and Young characterize as a bias in favour of observational adequacy. Fourth (C), on the basis of motivational factors, the delusion remains relatively unelaborated or circumscribed.

6.2 The Cotard Delusion

Stone and Young show how their style of account can be extended to the Cotard delusion in which a patient believes that she is emotionally dead, or

³⁶ It might be objected that this suggestion about the role of motivational factors like those that operate in more familiar cases of cognitive dissonance makes the psychological state of the delusional patient seem altogether too normal. On our view, the circumscription of the delusion may be intelligibly motivated to the extent that following through the consequences of the delusion would lead to substantial doxastic disruption and perhaps, in the limit, to the fracturing of the patient's conception of the world and his place in it. But this account of the fourth step does not involve us in underdescribing the strangeness of the patient's psychological state. Rather, it highlights the pathology of the third step. What is strange is the adoption and maintenance of the delusional hypothesis in the first place.

that she is not a real person, or simply that she is dead. It is proposed, first (E), that as a result of brain injury the patient either has a similar disorder of face processing to the one that figures in the Capgras delusion or else suffers from a more general flattening of affective responses to stimuli (Ramachandran and Blakeslee, 1998, p. 167; Gerrans, this volume; Young, this volume). Second (H), a hypothesis is generated. In its simplest and most dramatic form, the hypothesis is that the patient is dead. Corresponding to the two accounts of the anomalous experience in the first step, there are different possible accounts of how this delusional hypothesis comes to be prioritized as a candidate for belief.

On one account (given by Stone and Young), the Cotard patient's experience of other people's faces is the same as a Capgras patient's but because of a different attributional bias (associated with depression rather than persecution) a different hypothesis is generated to explain the nature of this experience. The patient locates the cause of the anomaly in herself ('I am dead') rather than in the external world ('My loved ones have been replaced by impostors').

On the other account, the experience of the Cotard patient when she sees the face of a loved one is different from that of a Capgras patient. This may seem to open up the option of saying that the two patients generate different hypotheses just because they accept different experiences of faces as veridical. We have already allowed that perhaps it seems to the Capgras patient that this person who looks just like the loved one is not really the loved one. But it is not plausible that we would give a correct account of the representational character of the Cotard patient's experience when looking at the face of a loved one by saying that it seems to the patient that she herself is emotionally dead. Nevertheless, it may be that the Cotard patient's hypothesis arises as an attempt to accept as veridical experiences of globally flattened affect, in which it seems to the patient that nothing has any significance for her and even that she is disembodied (Gerrans, this volume). From an initial delusional hypothesis, the patient might take one or two steps of non-conservative elaboration to arrive at the claim that she is not a real person or that she is dead.

The account, third (B), of how the hypothesis is adopted and maintained as a belief and, fourth (C), of why the delusion remains relatively circumscribed is as before.

6.3 The Fregoli Delusion

Patients suffering from the Fregoli delusion believe that they keep seeing someone close to them, perhaps their mother or brother, wherever they go or that they are being followed around by people who are known to them but who are unrecognizable because they are in disguise. In such a case, it is quite plausible to propose, first (E), that the patient has a heightened affective response even to unfamiliar faces (Ramachandran and Blakeslee, 1998, p. 171). When he sees someone whom he does not know and therefore cannot recognize he nevertheless experiences a sense of familiarity. Second (H), a hypothesis

is generated: 'These are people whom I know but they are in disguise.' There are two possible accounts of how this happens.

On one account, the hypothesis is generated to explain the patient's unusual experience. The hypothesis that they are people known to the patient explains the affective response; the hypothesis that they are in disguise explains the fact that the patient does not recognize them. This candidate is then prioritized over other explanatory hypotheses by an attributional bias. On the other account, the representational character of the patient's experience when seeing someone unfamiliar is correctly described by saying that it seems to the patient that this is someone whom he knows but does not recognize. He accepts these experiences as veridical and then slightly elaborates the delusion into the claim that he does not recognize these people because they are in disguise.

The account (B and C) of how the hypothesis is adopted and maintained as a belief and of why the delusion is not further elaborated is as before.

6.4 Mirrored-Self Misidentification and Mirror Agnosia

Some patients believe that the person they see in the mirror is not really them. But there seem to be rather different ways in which the same delusional belief may arise. In principle, mirrored-self misidentification might perhaps be explained along similar lines to the Capgras delusion. However, mirrored-self misidentification can occur without the suspiciousness that is characteristic of Capgras patients. One case reported in detail by Nora Breen and her colleagues (Breen et al., this volume) involves a patient (FE) with apparent right hemisphere damage and a problem with face recognition, prosopagnosia. FE believes that the person he sees in the mirror is not him but someone else who looks very much like him. A possible account of FE's case would be along the following lines. First (E), as a result of his brain damage, FE has anomalous experiences of familiar faces. When he looks in the mirror, the face that he sees looks quite like his but not exactly like his as he remembers it. Second (H), an obvious hypothesis is generated: The person in the mirror is someone else, not him. Then the third (B) and fourth (C) steps in the account – how the hypothesis is adopted and maintained as a belief and why the delusion is not further elaborated – would be as in the first three examples.

Breen and her colleagues also report a second case of mirrored-self misidentification, but in this case the patient (TH) has intact face processing. TH's problem is, rather, a loss of his understanding of mirrors (mirror agnosia). A possible account of TH's case would be along the following lines. First, when TH witnesses events in a mirror it seems to him that those events are happening in a separate location from the location of the events that are going on in front of the mirror. So when he sees himself in the mirror, it seems to him that this is someone who looks just like him but is in a different location from him. Second, accepting his perceptual experience as veridical he supposes that the person that he is seeing is in a different location from him and con-

cludes that the person cannot be him. The account would then continue along familiar lines.

6.5 Reduplicative Paramnesia

Patients suffering from reduplicative paramnesia believe that there are doubles of known people or places. For example, patient FE came to believe that his wife, Alison, was (so to speak) two people. One Alison was indeed his wife while 'the other Alison' was the mother of his children.³⁷

In another case that is described by Breen and her colleagues (this volume), a patient (DB) who knew that her husband had been dead for several years came to believe, following a stroke in the right parietal lobe, that her husband was nevertheless in the same hospital where she was being treated.³⁸ The medical condition of this patient was complex and she had a further delusional belief to the effect that she was able to use her left arm although in reality it was paralysed following the stroke. But in a speculative spirit we suggest that the reduplicative paramnesia for her husband might be explained as being the result, in part, of an anomalous experience of unfamiliar faces similar to that of a Fregoli patient. We suggest that, because of a deficit in DB's belief revision system, the delusional hypothesis that her husband was present in the hospital was not rejected. It was adopted and maintained as a belief even though DB also believed (correctly) that her husband was dead and despite the fact that she could recount the events surrounding his death and their consequences. We suggest, finally, that as a result of motivational factors DB's delusion about her husband remained relatively unelaborated.

6.6 Unilateral Neglect

Patients who suffer from left unilateral neglect, often following a stroke to the right parietal lobe, are indifferent to objects and events in the left side of their world, including sometimes the left side of their own bodies. A unilateral neglect patient may believe that his left limbs are not his own. If his left arm

³⁷ This is not, however, a straightforward case of duplication as FE believed that 'the other Alison' was not his wife Alison but could be mistaken for her. This is something like the idea that 'the other Alison' was an impostor.

³⁸ Someone might query whether this should really be classified as a case of reduplicative paramnesia; it might be said that believing that a husband who is long dead is also present and alive is not quite the same as believing in duplicates. In fact, the classification of cases is not crucial for our purposes. But we accept the classification of this case as reduplicative paramnesia by analogy with an imagined case in which the patient knows that her husband migrated to Canada long ago and is currently living there but also believes that her husband is present in the hospital in Sydney. It is worth noting that DB did have just such a pair of beliefs about her daughter. She knew that her daughter lived a considerable distance away but also believed that her daughter worked each day in the hospital kitchen, even while acknowledging that this combination of employment and domestic arrangements would be impossible for any one person.

is pricked with a pin, for example, then he may believe that it is actually his right arm that has been pricked. A neglect patient may even believe that his left limbs belong to someone else. In a memorable example reported by Eduardo Bisiach (1988, p. 469), a patient who insisted that his hand was the examiner's responded to the question, 'Ever see a man with *three* hands?' with: 'A hand is the extremity of an arm. Since you have three arms it follows that you must have three hands.' This belief is a case of monothematic and circumscribed delusion following brain injury.

It seems plausible that part of the basis of a patient's belief that his left arm is not really his is the paralysis, and loss of kinaesthetic and proprioceptive experience of the arm, immediately following his stroke. But (as we noted in section 2) many patients suffer from paralysis of the left arm after a stroke without becoming deluded. So there must be some additional factor at work. Unilateral neglect is not well understood and we shall not be offering any account of the relationship between neglect and the delusion. But it does seem that our four-step schematic account can be applied. We might assume, first (E), that as a result of damage to the right side of the brain the patient has anomalous experiences of the paralysed left side of his body. Second (H), a hypothesis is generated: His left limbs do not exist or are not really his. The hypothesis might be a prioritized attempt at explaining this strange experience or it might just be the result of accepting anomalous bodily experiences as veridical. Third (B), because of a deficit in the belief revision system the hypothesis is retained rather than being rejected, despite the fact that it conflicts with many of the patient's other beliefs, such as the belief that people do not usually go around with other people's arms attached to them. Fourth (C), for reasons given already, the delusion remains relatively unelaborated.

7. Extending the Account: Schizophrenic Delusions

Delusions are one of the characteristic symptoms of schizophrenia.³⁹ Some of the delusions suffered by schizophrenic patients are extreme cases of florid delusions; some are delusions that we have already considered such as the Capgras and Cotard delusions. But two delusions that are particularly associated with schizophrenia and to which our account might apply are the alien control delusion and the thought insertion delusion (the last two on the list of eight at the beginning). In the delusion of alien control, a patient believes that someone else is able to control his actions. In the delusion of thought insertion, a patient believes that someone else's thoughts are being inserted into her mind.

These delusions are among the so-called positive symptoms of schizo-

³⁹ *DSM-IV* (APA, 1994, pp. 285–6) requires two or more characteristic (Criterion A) symptoms for a diagnosis of schizophrenia, but allows that 'Only one Criterion A symptom is required if delusions are bizarre'. There are other criteria having to do with social or occupational dysfunction, with duration, and with exclusion of other possible diagnoses.

phrenia, along with other delusions, such as persecutory delusions, and hallucinations, particularly auditory hallucinations. The delusions of alien control and of thought insertion are sometimes called 'passivity phenomena' because they involve the denial of one's own agency (alien control) or initiation (thought insertion). If these delusions are to be brought within the scope of our schematic account then the key step is the first. We need to identify an anomalous experience that might prompt the generation of the delusional hypothesis. For the most part, the anomalous experiences that we have considered up to this point have been experiences resulting from perceptual or affective deficits. Now we need to allow for anomalies in our experience of intentional action and thought itself.

One of the most important ideas in recent research on schizophrenia has been that aspects of our experience of thought and agency may be underpinned by mechanisms of self-monitoring and particularly by comparisons that involve efference copies (Feinberg, 1978; Frith, 1992, ch. 5). We are all familiar with the idea that the brain makes use of a copy of the motor instruction sent to the muscles of the eye in order to make a prediction about the way in which the retinal array will change. Information about the actual change in the retinal array is then compared with this prediction. Where the actual change in the retinal array is as predicted it is 'cancelled' and the visual image remains stable. But, where there is a mismatch between the actual change and the change predicted on the basis of the motor instruction, this is interpreted visually as a shift in the world. This kind of shift is experienced if an eye movement is produced in a way that does not involve motor instructions to the muscles of the eye or if a motor instruction is issued but the eye does not move because of paralysis. In the absence of external manipulation of the eye or of paralysis, these comparisons based on efference copies permit discrimination between changes in us (an eye movement to the right) and changes in the world (a shift to the left).⁴⁰

Copies of motor instructions, efference copies, play other roles as well. For example, comparison between the position of a limb as predicted on the basis of an efference copy and the actual position of the limb according to visual or proprioceptive feedback can be used in the development of motor skills. Information about the difference between predicted position and actual position can be used to update the mapping on which the prediction rests, namely, the mapping between target limb positions and motor instructions.⁴¹ Once

⁴⁰ For a great deal more detail about eye position signals, see e.g. Jeannerod, 1988, ch. 4. For the notion of efference copy, see von Holst and Mittelstaedt, 1950; for the virtually identical notion of corollary discharge, see Sperry, 1950; for an authoritative review, see Jeannerod, 1997, ch. 6.

⁴¹ The motor instructions are determined, not just by the target limb positions, but by the target positions together with the current positions. Also, it is useful, here, to distinguish between the generation of motor instructions conceived as codes and the actual initiation of muscle movements achieved by sending the instructions from primary motor cortex to

that mapping has been learned, comparison between predicted limb position and target limb position allows revision of motor instructions in the light of changes in the position of a target object, without any need to wait for feedback information. Also, comparison between predicted limb position (according to the efference copy) and actual limb position (according to feedback information) can again serve as the basis for discrimination between changes that result from us and changes that result from the world.

Nielsen (1963) describes some experiments in which subjects are provided with misleading visual feedback about the movement of their hand. Normal subjects reach through a hole in a box and, using a pencil held in a gloved hand, try to follow a straight line that is printed on a sheet of paper. Subjects do not know that on some trials the hand that they see is not their own hand but the similarly gloved hand of the experimenter's assistant reflected in a mirror. The crucial trials are those in which the assistant traces about half way along the line and then starts deviating towards the right (as seen by the subject). On such trials, subjects typically produce a line (which they do not see, of course) deviating to the left. Of particular interest for our present purposes are subjects' reports of their own experiences under these experimental conditions. While the assistant's hand traced along the line, making the same movement as the subject's hand, the subject had a normal experience of controlling his or her hand movement. But, as the assistant's hand deviated from the straight line, this experience changed (1963, pp. 228–9):

On several occasions five [out of twenty] subjects spontaneously described that they directly felt that something outside themselves was pushing their hand towards the right or was resisting the free mobility of their hand. Ten other subjects declared on some occasions that they felt that their hand was sliding towards the right in a passive way.

Nielsen goes on to report how these subjects tried to explain their anomalous experience (1963, p. 229): '[The five subjects] suggested that "magnets", "unidentified forces", "invisible traces under the paper", or the like could be the cause.' Most subjects offered explanations in terms of environmental factors though '[some subjects] experienced right away that something was wrong with themselves and this alarmed them' (ibid.).

It seems that, in most of these subjects, a mismatch between hand position according to the prediction based on an efference copy and hand position according to visual feedback gave rise to an experience of something in the

the spinal cord and thence to the muscles. When we engage in motor imagery without actual muscle movement, motor instructions are apparently generated in pre-motor cortex and mapped to changes in limb position, even though no instructions are actually sent out to initiate movement. See Jeannerod, 1997, ch. 4. and Currie and Ravenscroft, 1997, for discussions of motor imagery.

environment robbing them of control of the movement of their hand. They reported that ‘they did not cause the bend to the right on purpose’ (*ibid.*).

7.1 Alien Control

In order to apply our four-step (EHBC) account of the aetiology of delusions to the delusion of alien control, we need to identify an anomalous experience that might prompt the generation of the delusional hypothesis. Comparisons involving efference copies could, of course, occur in an information-processing system from which consciousness was quite absent. But it does not seem absurd to suppose that our experience of our actions as initiated by ourselves depends in some way on information processing that involves efference copies.

The experimental results from Nielsen (1963) suggest that a cognitive abnormality, either in the generation of the efference copy or in the comparison between the efference copy and feedback information, might give rise to anomalous experiences of actions. The subject would be aware of trying to move his arm, say, in a particular way, but it would seem to the subject that the arm was being moved in a different way by some external force. We can also speculate that a more dramatic abnormality, in which no efference copy is even available for comparison, could give rise to anomalous experiences of a different kind. In this case, it might seem to the subject that his arm is moving, not of his own volition (since no efference copy is detected), but not exactly as the result of an external force either (since there is no definite mismatch between an efference copy and feedback information). It might seem to the subject that his arm is moving as the result of an agency that is not his own. It is fairly plausible that accepting experiences of one or other of these kinds as veridical would be a step on the way to the delusion of alien control.⁴²

In terms of our four-step schematic account, first (E), a cognitive abnormality related to efference copies gives rise to an anomalous experience, perhaps of an external force or an unknown agency, or perhaps simply of lack of the subject’s own volition. Second (H), a hypothesis is generated. The subject accepts the anomalous experience as veridical, and then elaborates the hypothesis that is thus prioritized by offering some more determinate account of the nature of the agency that is at work. Third (B), as the result of a deficit in

⁴² Spence et al., 1997, used neural imaging (PET) to investigate the neural basis of the experiences that give rise to the delusion of alien control. Seven schizophrenic patients who experienced passivity phenomena were scanned while they made voluntary movements using a joystick. Patients’ descriptions of their phenomenology during PET scanning included (p. 2001): ‘I felt like an automaton, guided by a female spirit who had entered me’; ‘I thought you [the experimenter] were varying the movements with your thoughts’; and ‘The spirits were moving my shoulder’. The brains of patients experiencing passivity phenomena showed hyperactivity in regions including the right inferior parietal lobule and the cingulate gyrus. Spence et al. note that other authors, including e.g. Cutting, 1989, have suggested the right parietal lobe as the possible origin of ‘alienation’ where this condition is characterized by qualities of detachment, ‘non-belongingness’ or spatial dislocation.

belief revision, the hypothesis is adopted and maintained as a belief despite its implausibility in the light of other things that the subject believes. Fourth (C), adoption and maintenance of the delusional hypothesis might be achieved without a major upheaval in the subject's belief system.

This fourth step figures in our account because we began from Stone and Young's suggestion about the aetiology of monothematic and circumscribed delusions. Stone and Young quite explicitly do not propose to apply their suggestion to more florid schizophrenic delusions. But it is consistent with the four-step account to allow that sometimes the motivational factors that support circumscription may fail. In such a case a subject may move in the direction of substantial and even disastrous doxastic disruption.

7.2 Some Other Symptoms of Schizophrenia

Before moving on to the delusion of thought insertion, we pause briefly to speculate on the possibility that abnormalities in the cognitive mechanisms implicated in comparisons involving efference copies may play a role in the aetiology of other schizophrenic symptoms.

Suppose, for example, that shifting the direction of one's attention can be regarded as analogous to moving one's arm. Then there would be room for two kinds of anomalous experiences involving attention. One would be an experience in which the subject's attention seems to be dragged somewhere by something external. The other would be an experience in which it seems to the subject that his attention shifts not of his own volition but not exactly as the result of an external attraction either. Perhaps it seems to shift under the control of a will other than his own. The first of these anomalous experiences might give rise, through a process of explanatory elaboration, to the hypothesis that there is something attention-claiming, such as people talking about the subject, going on at a particular location. It is possible that this kind of experience might play a role in the aetiology of delusions of reference in which the patient believes that other people are referring to him (Frith, 1987; Gold and Hohwy, this volume).

Suppose too that somewhat the same structure of comparisons is present in the mental activity of imagining hearing speech in one's own natural language. Someone might imagine, for example, hearing the critical voice of a parent. If an analogy between the activity of imagining hearing speech and the activity of moving one's arm can be made out at all then it might be that a cognitive abnormality could give rise to an impression of alien control of the words heard in one's mind. So it is conceivable that there may be the resources here to provide an outline account of certain anomalous experiences that are characteristic of schizophrenia, namely, auditory hallucinations (hearing voices) and of delusions that arise from those hallucinations.⁴³

⁴³ One complication here is that there is not literally any auditory feedback information to be compared with a prediction based on an efference copy. For a detailed account of auditory

7.3 Thought Insertion

Imagining hearing speech is different from imagining speaking, and inner speech can itself be of various kinds. Inner verbalized idle speculation is different from inner verbalized story telling and from inner verbalized serious thinking in pursuit of the truth, for example. Perhaps a cognitive abnormality related to efference copies could give rise to impressions of loss of control over one's inner speech. So it might seem to a subject that an external force, or an unknown agency, is causing him to experience inner speech. It might also be that a subject's impression of verbalized serious thinking that is directed by a will other than his own is different from an impression of verbalized story telling or verbalized idle speculation that is under alien control. An anomalous experience of the first kind might play a role in the aetiology of the delusion that verbalized thoughts are being inserted into the subject's mind.

Thinking itself is carried out, at least sometimes, as part of some specific intellectual project. I might set out to think through the consequences of some proposition P that I am inclined to believe, checking those consequences for plausibility in the light of a further battery of knowledge. As I carry out this project, information about the course of my actual thinking is, presumably, checked against my original intention. But then there is the possibility that my actual thinking might fail to measure up to my intention or indeed that there might be no record of that intention available for comparison. Given what has already been said about the experience of alien control, it is at least conceivable that this kind of mismatch or comparison failure would be experienced as serious thought about P and its consequences taking place in my mind yet without being intended or initiated by me.⁴⁴ An anomalous experience of that kind could play a role as the first step in the aetiology of the delusion of thought insertion; and this could be so quite independently of any equation of thought with inner speech.

8. Conclusion: Delusions and Theory of Mind

At the end of section 5, we outlined a four-step account of the aetiology of monothematic and circumscribed delusions following brain injury: Experience,

hallucinations that sets the operation of the 'mind's ear' and the 'mind's voice' against the background of a box-and-arrow model of normal speech perception and production, see David, 1994.

⁴⁴ This is intended to be in the spirit of the proposal by Frith, 1992, ch. 5, that the experience of thought insertion arises from a failure of central monitoring. In fact, Frith's account involves two kinds of central monitoring. There is the monitoring of actions, which enables us to distinguish 'changes due to our own actions and changes due to external events' (1992, p. 81); and there is also the monitoring of intentions, which enables us to distinguish 'between actions caused by our own goals and plans (willed actions) and actions that are in response to external events (stimulus-driven actions)' (*ibid.*). For discussion of Frith's application of these ideas about central monitoring to the case of thought, see Currie, this volume; Gold and Hohwy, this volume; and Campbell, 1999.

Hypothesis, Belief, Circumscription. In that account, the third step is that a hypothesis is adopted and maintained as a belief despite the fact that it is inconsistent with many other things that the subject believes. We suggested, cautiously and provisionally, that this resistance against conflicting considerations might be explained in terms of impairment to, or even abolition of, some component of the cognitive system responsible for belief revision and that this might be a further result of brain injury.⁴⁵ In section 6, we applied the EHBC account to examples of frankly neuropsychological delusions and, in section 7, we showed how it might plausibly be extended to certain delusions that occur in schizophrenic patients. The extension of the account to schizophrenic delusions, if it is taken to include the cautious suggestion, involves the claim that at least some patients with schizophrenia have a deficit in the cognitive mechanisms that underpin belief revision.

We have already noted that there is an association between delusions and right hemisphere damage. Furthermore, delusions occur in some patients with unilateral neglect and these patients may have intact left hemispheres. So it seems quite likely that, if there is indeed a belief revision deficit that figures in the aetiology of delusions, then this deficit normally results from damage to the right hemisphere. If the extension of the account to schizophrenic delusions is correct, then in at least some cases of schizophrenia there must be impairment to cognitive functions that are, in normal subjects, subserved by right hemisphere neural mechanisms.

Robyn Langdon and her colleagues (1997) used a picture-sequencing task to assess theory of mind abilities in twenty schizophrenic patients.⁴⁶ A subgroup of six patients was found to have a specific problem with the task in those cases where the story involved a character acting on a false belief.⁴⁷ Also, in a study of non-clinical adults, Langdon and Coltheart (1999) found that high schizotypal subjects showed specific theory of mind problems that could not be accounted for in terms of either executive planning deficits or failure to inhibit cognitively salient but inappropriate information.⁴⁸ Because delusions

⁴⁵ See Langdon and Coltheart, this volume, for a more forthright commitment to this proposal.

⁴⁶ We speak of 'theory of mind' abilities without prejudging the question whether the basis of those abilities is a matter of knowing a psychological theory, or being able to engage in mental simulation, or something else.

⁴⁷ These patients were able to put the pictures in the correct order when the stories involved causal relations or social scripts, for example. The remaining fourteen patients either succeeded for all categories of story or else had a general sequencing difficulty and no additional difficulty with false belief stories.

⁴⁸ Experimental work with schizophrenic patients is complicated by the fact that the subjects are typically receiving medication and these antipsychotic drugs themselves have cognitive effects. See Williams et al., 1998, for discussion. However, when the same pattern of task performance is found in patients who are receiving medication and in non-clinical high schizotypal subjects (who are not receiving medication) this supports the hypothesis that the patients' performance is not just a result of their medication.

are one of the characteristic symptoms of schizophrenia,⁴⁹ these findings are at least suggestive of an association between delusions and theory of mind problems in schizophrenic patients.

Langdon and her colleagues (1997) did not, in fact, find any correlation between reality distortion symptoms such as delusions and hallucinations, on the one hand, and membership of the subgroup of schizophrenic patients with specific theory of mind problems, on the other.⁵⁰ But that lack of correlation is consistent with the speculation that there is a common factor in the aetiologies of both delusions and theory of mind problems. Perhaps there is some cognitive function that is implicated in normal belief revision and also in the understanding of false belief stories.

If there is a cognitive function whose impairment is a factor in the aetiology of delusions then, we have already said, it is plausible that it is normally subserved by right hemisphere neural mechanisms. So it is of some interest that Francesca Happé and her colleagues (1999) report theory of mind problems in patients who have suffered right hemisphere strokes.

In this study, fourteen right hemisphere stroke patients (average age 64 years) and a group of healthy elderly control subjects were given short passages to read, with a question after each passage. When the passage was a theory of mind story, giving a correct answer to the question involved making an inference about the thoughts, intentions or feelings of a character in the story. When the passage was a non-mental story, answering the question correctly involved making an inference about, for example, physical causation. The right hemisphere stroke patients scored less well than controls on the theory of mind stories but not on the non-mental stories. Also, the right hemisphere patients scored less well on theory of mind stories than on non-mental stories, whereas this was not so for control subjects. In addition, five left hemisphere stroke patients were tested on modified versions of the tasks and had no greater difficulty with the theory of mind stories than with the non-mental stories.⁵¹

These results are consistent with the idea that some cognitive function that is required for success on theory of mind tasks is normally subserved by right hemisphere neural mechanisms. It is also consistent, therefore, with the speculation that there is some cognitive function, normally subserved by right hemi-

⁴⁹ The presence of delusions is virtually sufficient for a diagnosis of schizophrenia if the delusions are of the bizarre variety (see again note 39).

⁵⁰ However, Doody et al., 1998, found significant correlations between performance on (relatively difficult) theory of mind tasks and ratings of positive symptoms such as delusions and hallucinations. Also, Drury, Robinson and Birchwood, 1998, found poor theory of mind performance when schizophrenic patients were in acute episodes (characterized by delusions, hallucinations and bizarre speech or behaviour) but not when the patients had recovered from acute episodes.

⁵¹ The tasks had to be modified because left hemisphere patients typically have impaired language.

sphere structures, that is crucial both for normal belief revision and for theory of mind abilities.⁵²

This is just a speculation and the reality may be much less interesting. It may be, for example, that there are mechanisms that are implicated in belief revision and functionally quite distinct mechanisms that are involved in theory of mind abilities. Perhaps these mechanisms are neuroanatomically close to each other in the right hemisphere so that right hemisphere damage or dysfunction sometimes leads to co-occurrence of delusions and theory of mind problems. Even if there is a cognitive common factor in some cases of these co-occurring symptoms, the line of thought that we have been exploring is complicated by the fact that we cannot assume that theory of mind problems have the same basis across different groups of subjects.

In this context, it is a striking fact that deaf children of hearing parents show performance on theory of mind tasks much like that of autistic children. Peterson and Siegal (this volume) propose that this is because deaf children of hearing parents have very limited opportunities for conversation with their parents. Because of the absence of a shared language between child and parents, topics of conversation are restricted to things that are perceptibly present and there is little discussion of more abstract matters such as beliefs, emotions and reasons for action. This means that deaf children of hearing parents are deprived of the cognitive input that would normally play a role in the development of everyday psychological understanding and theory of mind abilities.⁵³

Despite these complications, it may be that there is a cognitive common factor involved in both some cases of delusions and some cases of theory of mind problems. So it is important to consider how alternative theories of everyday psychological understanding (including understanding of false beliefs) might account for the co-occurrence of those symptoms. Frith (1992) suggests that both delusions and theory of mind problems flow from impairments to the mechanisms of meta-representation (Leslie, 1987, 1994). Currie (this volume) criticizes this suggestion and proposes that delusions may arise from a disorder of imagination in which what is imagined is mistaken for what is believed. According to the simulation theory of everyday psychological understanding,

⁵² The stroke patients investigated by Happé and her colleagues were not delusional. So if there is some right hemisphere cognitive function that is implicated in normal belief revision and also in the understanding of false belief stories then impairment of that function is not sufficient, by itself, for the production of delusions.

⁵³ Peterson and Siegal, this volume, also observe that the situation is quite different in the case of deaf children who grow up with signing deaf parents. These children, who are 'native signers', develop theory of mind abilities at the same age as children with normal hearing. Paul Harris proposes, 1996, p. 211, 'that children learn about thoughts and beliefs in the context of conversation' and notes the prediction that 'children with limited or delayed exposure to conversation (e.g. deaf children) should show difficulties on tests of belief understanding' (1996, p. 220).

of course, imagination would also be a vital factor in successful performance on false belief tasks.

The occurrence of theory of mind problems in patients with a psychiatric diagnosis characterized by delusions may turn out to pose an important challenge for theories of everyday psychological understanding. But in order to judge the success of competing theories in measuring up to the challenge we first need to investigate the nature of delusions themselves. Our aim has been to offer a framework for that investigation.

Corpus Christi College
University of Oxford

Macquarie Centre for Cognitive Science
Macquarie University

References

- American Psychiatric Association 1994: *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV). Washington, DC: American Psychiatric Association.
- Benson, D.F., Gardner, H. and Meadows, J.C. 1976: Reduplicative paramnesia. *Neurology*, 26, 147–51.
- Bentall, R.P. 1994: Cognitive biases and abnormal beliefs: Towards a model of persecutory delusions. In A.S. David and J.C. Cutting (eds), *The Neuropsychology of Schizophrenia*. Hove, E. Sussex: Psychology Press, 337–60.
- Bentall, R.P. and Kinderman, P. 1998: Psychological processes and delusional beliefs: Implications for the treatment of paranoid states. In T. Wykes, N. Tarrier and S. Lewis, *Outcome and Innovation in the Psychological Treatment of Schizophrenia*. Chichester: John Wiley and Sons, 119–44.
- Bermúdez, J.L. Forthcoming: Normativity and rationality in psychiatric disorders.
- Berrios, G.E. 1991: Delusions as ‘wrong beliefs’: A conceptual history. *British Journal of Psychiatry*, 159, 6–13.
- Berrios, G.E. and Luque, R. 1995: Cotard’s syndrome: Analysis of 100 cases. *Acta Psychiatrica Scandinavica*, 91, 185–8.
- Bisiach, E. 1988: Language without thought. In L. Weiskrantz (ed.), *Thought Without Language*. Oxford University Press, 464–84.
- Campbell, J. 1999: Schizophrenia, the space of reasons, and thinking as a motor process. *The Monist*, 82, 609–25.
- Capgras, J. and Carrette, P. 1924: Illusion des sosies et complexe d’Oedipe. *Annales Médico-Psychologiques*, 12, 48–68.
- Capgras, J. and Reboul-Lachaux, J. 1923: L’illusion des ‘sosies’ dans un délire systématisé chronique. *Bulletin de la Société Clinique de Médecine Mentale*, 11, 6–16.
- Cherniak, C. 1986: *Minimal Rationality*. Cambridge, MA: MIT Press.

- Church, J. 1987: Reasonable irrationality. *Mind*, 96, 354–66.
- Cotard, J. 1882: Du délire des négations. *Archives de Neurologie*, 4, 152–70, 282–95.
- Courbon, P. and Fail, G. 1927: Syndrome d'illusion de Fregoli' et schizophrénie. *Bulletin de la Société Clinique de Médecine Mentale*, 20, 121–5
- Currie, G. and Ravenscroft, I. 1997: Mental simulation and motor imagery. *Philosophy of Science*, 64, 161–80.
- Cutting, J. 1989: Body image disorders: Comparison between unilateral hemisphere damage and schizophrenia. *Behavioural Neurology*, 2, 210–10.
- David, A.S. 1994: The neuropsychological origin of auditory hallucinations. In A.S. David and J.C. Cutting (eds), *The Neuropsychology of Schizophrenia*. Hove, E. Sussex: Psychology Press, 269–313.
- Davidson, D. 1984: *Inquiries into Truth and Interpretation*. Oxford University Press.
- de Pauw, K.W., Szulecka, T.K. and Poltock, T.L. 1987: Fregoli syndrome after cerebral infarction. *Journal of Nervous and Mental Diseases*, 175, 433–8.
- Dennett, D.C. 1981: Brain writing and mind reading. In *Brainstorms: Philosophical Essays on Mind and Psychology*. Brighton: Harvester Press, 39–50.
- Dennett, D.C. 1987: *The Intentional Stance*. Cambridge, MA: MIT Press.
- Doody, G.A., Gotz, E.C., Johnstone, E.C., Frith, C.D. and Cunningham Owens, D.G. 1998: Theory of mind and psychosis. *Psychological Medicine*, 28, 397–405.
- Drury, V.M., Robinson, E.J. and Birchwood, M. 1998: 'Theory of mind' skills during an acute episode of psychosis and following recovery. *Psychological Medicine*, 28, 1101–12.
- Eilan, N. 2000: On understanding schizophrenia. In D. Zahavi and J. Parnas (eds), *Exploring the Self: Philosophical and Psychopathological Perspectives on Self Experience*. Amsterdam: John Benjamins.
- Ellis, H.D. and de Pauw, K.W. 1994: The cognitive neuropsychiatric origins of the Capgras delusion. In A.S. David and J.C. Cutting (eds), *The Neuropsychology of Schizophrenia*. Hove, E. Sussex: Psychology Press, 317–35.
- Ellis, H.D., Whitley, J. and Luauté, J.P. 1994: Delusional misidentification: The three original papers on the Capgras, Fregoli and intermetamorphosis delusions. *History of Psychiatry*, 5, 117–46.
- Ellis, H.D. and Young, A.W. 1990: Accounting for delusional misidentifications. *British Journal of Psychiatry*, 157, 239–48.
- Ellis, H.D., Young, A.W., Quayle, A.H. and de Pauw, K.W. 1997: Reduced autonomic responses to faces in Capgras delusion. *Proceedings of the Royal Society: Biological Sciences*, B264, 1085–92.
- Enoch, M.D. and Trethowan, W.H. 1991: *Uncommon Psychiatric Syndromes*, 3rd edn. Oxford: Butterworth-Heinemann.
- Feinberg, I. 1978: Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin*, 4, 636–40.
- Festinger, L. 1957: *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Fodor, J. 1987: *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

- Foley, J.M. and Breslau, L. 1982: A new syndrome of delusional misidentification. *Annals of Neurology*, 12, 76.
- Förstl, H., Almeida, O.P., Owen, A.M., Burns, A. and Howard, R. 1991: Psychiatric, neurological and medical aspects of misidentification syndromes: A review of 260 cases. *Psychological Medicine*, 21, 905–10.
- Frith, C.D. 1987: The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine*, 17, 631–48.
- Frith, C.D. 1992: *The Cognitive Neuropsychology of Schizophrenia*. Hove, E. Sussex: Lawrence Erlbaum Associates.
- Garety, P.A. and Hemsley, D.R. 1994: *Delusions: Investigations into the Psychology of Delusional Reasoning*. Oxford University Press. (Reprinted, Hove, E. Sussex: Psychology Press, 1997.)
- 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.
- Gendler, T.S. 2000: The puzzle of imaginative resistance. *Journal of Philosophy*, 97, 55–81.
- Happé, F., Brownell, H. and Winner, E. 1999: Acquired 'theory of mind' impairments following stroke. *Cognition*, 70, 211–40.
- Harris, P.L. 1996: Desires, beliefs, and language. In P. Carruthers and P.K. Smith (eds), *Theories of Theories of Mind*. Cambridge University Press, 200–20.
- Higgins, E.T. 1987: Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94, 319–40.
- Hirstein, W. and Ramachandran, V.S. 1997: Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons. *Proceedings of the Royal Society: Biological Sciences*, B264, 437–44.
- 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, 801–12.
- Jaspers, K. 1963: *General Psychopathology*, trans. J. Hoenig and M.W. Hamilton. Manchester: Manchester University Press.
- Jeannerod, M. 1988: *The Neural and Behavioural Organization of Goal-Directed Movements*. Oxford University Press.
- Jeannerod, M. 1997: *The Cognitive Neuroscience of Action*. Oxford: Blackwell Publishers.
- Kaney, S. and Bentall, R.P. 1989: Persecutory delusions and attributional style. *British Journal of Medical Psychology*, 62, 191–8.
- Kaney, S. and Bentall, R.P. 1992: Persecutory delusions and the self-serving bias. *Journal of Nervous and Mental Disease*, 180, 773–80.
- Kinderman, P. 1994: Attentional bias, persecutory delusions and the self concept. *British Journal of Medical Psychology*, 67, 53–66.
- Langdon, R. and Coltheart, M. 1999: Mentalising, schizotypy, and schizophrenia. *Cognition*, 71, 43–71.
- Langdon, R., Michie, P.T., Ward, P.B., McConaghy, N., Catts, S. and Coltheart, M. 1997: Defective self and/or other mentalizing in schizophrenia: A cognitive neuropsychological approach. *Cognitive Neuropsychiatry*, 2, 167–93.

- Leslie, A.M. 1987: Pretense and representation: The origins of 'theory of mind'. *Psychological Review*, 94, 412–26.
- Leslie, A.M. 1994: Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50, 211–38.
- Lewis, D. 1986: *On the Plurality of Worlds*. Oxford University Press.
- Lyon, H.M., Kaney, S. and Bentall, R.P. 1994: The defensive function of persecutory delusions: Evidence from attribution tasks. *British Journal of Psychiatry*, 164, 637–46.
- Mackie, J., Ebmeier, K.P. and O'Carroll, R.E. 1994: An MRI, SPECT and neuropsychological study of a patient presenting with Capgras syndrome. *Behavioural Neurology*, 7, 211–15.
- 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 T.F. Oltmanns and B.A. Maher (eds), *Delusional Beliefs*. Chichester: John Wiley and Sons, 15–33.
- Maher, B.A. 1992: Delusions: Contemporary etiological hypotheses. *Psychiatric Annals*, 22, 260–8.
- Nielsen, T.I. 1963: Volition: A new experimental approach. *Scandinavian Journal of Psychology*, 4, 225–30.
- Ramachandran, V.S. and Blakeslee, S. 1998: *Phantoms in the Brain: Human Nature and the Architecture of the Mind*. London: Fourth Estate.
- Spence, S.A., Brooks, D.J., Hirsch, S.R., Liddle, P.F., Meehan, J. and Grasby, P.M. 1997: A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain*, 120, 1997–2011.
- Sperry, R.W. 1950: Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, 43, 482–9.
- Stone, T. and Young, A.W. 1997: Delusions and brain injury: The philosophy and psychology of belief. *Mind and Language*, 12, 327–64.
- Tranel, D., Damasio, H. and Damasio, A.R. 1995: Double dissociation between overt and covert recognition. *Journal of Cognitive Neuroscience*, 7, 425–32.
- von Holst, E. and Mittelstaedt, H. 1950: Das Refferenzprinzip: Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften*, 37, 464–76.
- Williams, J.H., Wellman, N.A., Geaney, D.P., Cowen, P.J., Feldon, J. and Rawlins, J.N.P. 1998: Reduced latent inhibition in people with schizophrenia: An effect of psychosis or of its treatment. *British Journal of Psychiatry*, 172, 243–9.