

Rationality and Schizophrenic Delusion

IAN GOLD AND JAKOB HOHWY

Abstract: The theory of rationality has traditionally been concerned with the investigation of the norms of rational thought and behaviour, and with the reasoning procedures that satisfy them. As a consequence, the investigation of irrationality has largely been restricted to the behaviour or thought that violates these norms. There are, however, other forms of irrationality. Here we propose that the delusions that occur in schizophrenia constitute a paradigm of irrationality. We examine a leading theory of schizophrenic delusion and propose that some delusions can be traced to a violation of a condition on thought we call *egocentricity*. We argue that the violation of egocentricity leads to irrational states that cannot be explained by the traditional categories of irrationality and conclude, therefore, that these states belong in a new branch of the theory of irrationality, that of *experiential irrationality*.

1. Introduction

The investigation of rationality is both normative and empirical. It is normative because it deals with the standards according to which one *ought* to act if one is to count as rational; it is empirical because we take it to be a conceptual truth that rationality is a property of actual reasoning and behaving agents, and a theory of rationality has to be a theory that some such agents do in fact satisfy. The study of rationality is thus one of the many places where a priori philosophy meets psychology and the other human sciences.

A good deal of the theory of rationality has been devoted to conceptual, rather than empirical, investigation. It has focused on rules of various kinds—those of logic, decision-theory, rational choice theory, and the like—in an effort to articulate the ideals of rational thought and behaviour. More recently, however, the relevance of empirical psychology, particularly the psychological theory of reasoning, has been recognized as significant (see Harman, 1986). We follow this development by arguing that the study of psychopathology is also relevant to the theory of rationality. We argue that the case of schizophrenic delusion does not fit into the traditional branches of the theory of rationality and propose, therefore, what Lewis (1986) calls a new ‘department’ of rationality.

It is not our intention to develop a complete theory of delusion in schizo-

We are very grateful to Max Coltheart, Martin Davies, Joel Gold and Natalie Stoljar for many helpful comments on earlier drafts of this paper.

Address for correspondence: Ian Gold, Vision Research Unit, McGill University, Royal Victoria Hospital, 687 avenue des Pins ouest, Montreal, Quebec, Canada H3A 1A1.

Email: ian@vision.mcgill.ca.

phrenia, much less a general theory of delusion (see Breen et al., and Langdon and Coltheart, both in this volume). Rather, we aim to offer an account of a subset of the delusions of schizophrenia and explore their implications for the study of rationality. We claim that schizophrenic delusion is a paradigm case of irrationality and, based on a discussion of a leading theory of delusion—that of Christopher Frith (1987, 1992)—we hypothesize that it is brought about by a violation of a constraint on rational thought we call *egocentricity*, a thought's manifesting itself in consciousness as having originated in one's own mind. Our account of schizophrenic delusion attempts to explain the empirical facts and to be consistent with what is known about schizophrenia. Our approach to the irrationality of delusion is thus naturalistic. We argue that neither of the traditional conceptions of rationality offers a successful account of schizophrenic delusion¹ and conclude that a new branch of the theory of rationality is required. Because this branch deals with the rationality of experience, we call it *experiential rationality*.

2. The Traditional Departments of Rationality

2.1 Procedural Rationality

The study of rationality since Hume has concerned itself primarily with modes of reasoning and its behavioural consequences. A wide variety of types of reasoning has been investigated, including instrumental reasoning, utility-maximization, and the like, but in all cases the focus has been on the properties of the reasoning rules, algorithms, or, as we will say, *procedures* (Brown, 1988). Some theories have explicated procedures as ideals, whereas others have attempted to be more psychologically realistic, but, according to the vast majority of these theories, rationality is a matter of rule-satisfaction. As Brown (1988) puts it:

Rules are at the heart of our classical model of rationality: if we have universally applicable rules, then all who begin from the same information must indeed arrive at the same conclusion (p. 19).²

¹ The psychological theory of delusion is obviously distinct from the theory of delusional irrationality. However, it will become clear below that the typical psychological accounts of delusion fall quite straightforwardly into the traditional branches of the theory of rationality. For this reason, we consider both the theory of the delusion and the theory of delusional irrationality together.

² Cf. also Rescher, 1988, who emphasizes that the rational is distinguished from individual, idiosyncratic or subjective desires and intentions: 'Above all, reason is systematic: it requires us to pursue intelligently adopted objectives in intelligent ways, acting on principles that make sense in a systematic way and whose appropriateness other agents can in principle also determine . . . Rationality demands: consistency, uniformity, coherence, simplicity, economy' (pp. 16–18).

On the procedural account, therefore, one is *irrational* if one's thoughts or actions fail to be governed by the relevant ideal, rule, or reasoning procedure.³ For example, a paradigm case of procedural irrationality is being susceptible to what philosophers call 'Dutch books' (Skyrms, 1986)—arguments that reveal one's behaviour to be logically equivalent to a betting strategy that guarantees financial loss.

Central to the procedural account is the claim that the particular contents of one's thoughts are *not* in general relevant to the analysis of rationality. As Hume (1888/1978) famously remarked 'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger' (p. 416). These contents are the contents of desires (which Hume called the 'passions') and beliefs, and they form no part of this domain of rationality: 'Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them' (p. 415).

2.2 Content Rationality

This is not to say, however, that the content of belief and desire is not a matter of rationality but only that it belongs to a different branch of its study.⁴ David Lewis (1986) puts this view as follows:

instrumental [i.e. procedural] rationality, though it is the department of rationality that has proved most tractable to systematic theory, remains only one department among others. We think that some sorts of belief and desire (or, of dispositions to believe and desire in response to evidence) would be unreasonable in a strong sense—not just unduly sceptical or rash or inequitable or dogmatic or wicked or one-sided or short-sighted, but utterly unintelligible and nonsensical. Think of the man who, for no special reason, expects unexamined emeralds to be grue.⁵ Think of Anscombe's (1957) example (in *Intention*, section 37) of someone with a basic desire for a saucer of mud. These beliefs and desires are unreasonable; though if twisted desire is combined with correspondingly twisted belief, then it may be that the failing lies entirely outside the purview of the department of instrumental ration-

³ For various approaches to the procedural account of rationality, see, for example, Cherniak, 1986; Kaplan, 1996; Rescher, 1988. For a discussion of rationality that takes psychological reality into account, see Kahneman, Slovic and Tversky, 1982.

⁴ We do not mean to suggest that the departments of procedural and content irrationality never intersect, or that there are never ambiguous cases. A familiar example of such an intersection is the case of a belief with a self-contradictory content. The content of the belief is irrational because it exemplifies a procedural violation. (We are grateful to Martin Davies for alerting us to this issue.)

⁵ 'Grue' is the gerrymandered property of being green if examined before some time, *t*, in the future and blue if examined thereafter (see Goodman, 1973). The gerrymandered quality of the property is what makes a belief that emeralds are grue ineligible.

ality. So I say that other departments of rationality also may have a constitutive role. What makes the perversely twisted assignment of content incorrect, however well it fits the subject's behavior, is exactly that it assigns ineligible, unreasonable content when a more eligible assignment would have fit behavior equally well (pp. 38–9).

We agree that there are kinds of irrationality that can be traced to the contents of particular beliefs and desires. We argue, however, that there are cases of irrationality that represent neither procedural failings nor beliefs or desires with ineligible content in Lewis's sense. These cases are delusions.

3. A Case of Irrationality: Delusion in Schizophrenia

3.1 Delusion

Delusion is one of the five characteristic symptoms identified by the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV, 1994) the presence of which can lead to a diagnosis of schizophrenia. Indeed, delusion is a pervasive and highly typical symptom of schizophrenia, though by no means unique to it (see Breen et al.; Gerrans; and Young, all in this volume). Some of the characteristic delusions of schizophrenia include *thought broadcasting*, the belief that one's thoughts are being heard by others; *thought insertion*, the belief that some of one's thoughts are being inserted into one's mind from an outside source; *delusions of control*, the belief that one is being controlled by an external force; *delusions of reference*, the belief that the actions of others have a special reference to oneself; and *delusions of persecution*, the belief that one is the target of the malicious actions of others (Cutting, 1990).

To an outside observer, delusions are, first and foremost, bizarre and wildly improbable:

Delusions of being wired or radio controlled are relatively common [in schizophrenia]. Often it is the FBI or the CIA which is the suspected perpetrator of the scheme. One patient was convinced that he had a radio sewn into his skull when he had had a minor scalp wound sutured and had tried to bring legal suit against the FBI innumerable times. Another man, at one time a highly successful superintendent of schools, became convinced that a radio had been implanted in his nose. He went to dozens of major medical centers, even to Europe, seeking a surgeon who would remove it. He even had an X-ray of his nose showing a tiny white speck which he was convinced was the radio (Torrey, 1995, p. 54).

From a theoretical point of view, however, what is more significant than the bizarreness of delusion is the fact that, in the delusional state, the patient's mind seems invaded by an alien force. The classic description of the phenomenology of delusion by Kraepelin (1919) conveys this well:

People speak to the patient in his thoughts, guide them, contradict him, 'offer' him thoughts, suggest them to him, transfer to him words, thoughts, pictures, smells and feelings. A patient said, 'My senses don't belong to me any more, they are being unlawfully taken from me.' Strangers send him thoughts silently and speak in his head, it is 'a remembrance, a memory a memorial,' a 'receiving of thoughts.' In this way his own thoughts are disturbed, 'drilled,' 'drawn off'; he cannot think when the voice speaks. A patient explained, 'They take my thoughts from me and nothing comes back but a ragamuffin.' What he thinks himself is distorted; his thoughts are 'plundered, organised and published.' 'The voices and my brain are one, I must think what the voice says,' said a patient, and a woman complained: 'The voices work on my thoughts from morning to evening, suggest dreams to me and torment me unceasingly.' Many of the patients must utter aloud their own thoughts or those that are given to them, 'low by movements of the lips,' 'say silly stuff to oneself.' 'It flows into the brain as a thought and expresses itself as words in the mouth,' said a patient. Another heard 'dead' and had to answer 'bread' (pp. 12–13).

The alien quality of delusional experience is, in our view, the key to understanding at least some of the delusions of schizophrenia and the nature of the irrationality of schizophrenic delusion.

That delusions *are* irrational we take to be obvious. Like Dutch book arguments and a desire for a saucer of mud, the delusions described above are *paradigms* of irrationality. Adopting a deficient betting strategy, as in the case of Dutch books, surely represents a misdemeanor of irrationality compared to the delusion that the FBI is controlling you by a radio surreptitiously sewn into your skull. We will assume, therefore, that a theory of rationality that does not explain why delusions such as these count as irrational is an inadequate theory. For the sake of simplicity we focus on the delusion of thought insertion which is widely thought to be an almost certain indication of schizophrenia (Torrey, 1995), but we suspect that other delusions will also be explained by this, or a similar, account.

3.2 The Etiology of Delusion

One of the central theories of the etiology of delusion in schizophrenia is that of Frith (1987, 1992), and we turn now to a consideration of his model.⁶ Our

⁶ For a review of other accounts, see Garety and Freeman, 1999.

proposal, however, does not depend on the adequacy of that model.⁷ We take it that the constraint on rationality we identify by exploring this model is a genuine constraint whether or not it is actually violated in schizophrenia, or violated in the way Frith suggests. However, we accept that should it be discovered that there are *no* psychological conditions in which this constraint is in fact violated, it would cease to be relevant to the theory of rationality. Our view is thus susceptible of empirical refutation.

Frith's account first distinguishes between two routes to the formation of an intention to act. The first is stimulus-driven; the second is driven by goals and plans. In the first case, a particular stimulus leads a subject to form the intention to act, as when, for example, one forms the intention to order chocolate cake after seeing chocolate cake on the menu. Frith calls these *stimulus-intentions*. In the second case, the formation of an intention to act does not follow upon a stimulus but upon a goal or plan. One might, for example, form the intention to get a French dictionary in order to pursue the goal of learning French. Frith calls these *willed intentions*. Frith further posits the existence of a cognitive monitor that keeps track of both kinds of intention as well as of the actions that are actually chosen by the subject. In effect, the monitor is a mechanism of *metarepresentation*: it represents the ordered pair of the intention (whether stimulus or willed) and the action chosen as a result of that intention (see Frith, 1992). Crucially, the effect of metarepresenting an intention is to bring that intention into the subject's consciousness.

Frith's hypothesis is that schizophrenic delusion is caused by *a failure of the monitor to represent willed intentions* (see also Feinberg, 1978; Feinberg and Guazzelli, 1999). Consider a simple example. It is one of my goals to learn French. At some stage, I form the willed intention to buy a French dictionary. The monitor metarepresents that intention and thereby brings it into consciousness. Having formed the intention to buy the dictionary, I initiate the action of buying it (all things being equal). The monitor metarepresents both my action and my intention and can thus verify that the action executed satisfies the relevant intention. Figure 1 gives a sketch of the view.

Now suppose that the process of self-monitoring is defective and the monitor fails to represent the intention to buy a dictionary. In this case, I would find myself buying a dictionary without any conscious awareness that I had formed the intention to do so, and I would therefore be faced with a puzzle: an action has been initiated, but I have no awareness of any intention to initiate it. I might, therefore, be tempted to look elsewhere for an explanation. One possibility is that some external force initiated the action for a reason of its own. This belief would of course be mistaken, indeed, delusional—a delusion of control.

Consider a different example from Frith (1987). I am in the tea-room talk-

⁷ For some problems with Frith's model, see Currie, this volume.

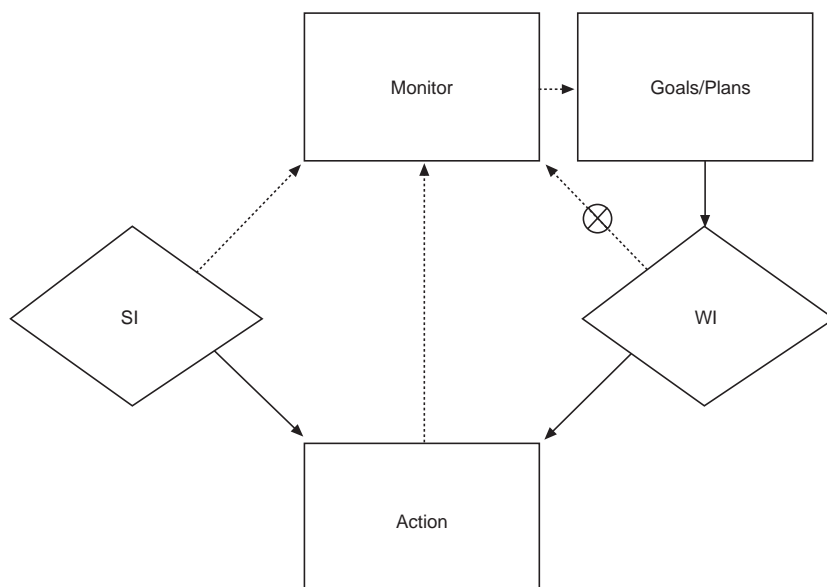


Figure 1 Frith's model of the failure of self-monitoring. If the action is a thought, it may be (a) perceived in the form of a hallucination; (b) experienced as having been caused by an external agency as in the delusion of thought insertion or the delusion of passivity; or (c) the result of an unassociated stimulus as in the delusion of reference. WI = willed intention; SI = stimulus intention; = feedback; ⊗ = lesion.

ing to Tim, and I overhear John talking about wine. Since it is one of my goals to learn about wine, I form the intention to switch my attention from Tim to John. However, if I cannot monitor my intentions, I will find myself listening to John without any awareness of an intention to do so, and I might suppose that this happened because John was talking about me, and that is what got my attention—a delusion of reference.

If one now interprets 'action' in the model (see Figure 1) as including the thinking of a thought (as Frith intends), it is possible to account for other of the experiential symptoms of schizophrenia (Frith, 1987, 1992). Suppose one of my goals is to go to work on time each morning. On a particular morning, I form the intention to catch the bus and then perform the action of thinking 'Catch the bus!' If I have not monitored the intention to catch the bus, I will find myself with a particular thought without any awareness of the intention that initiated it. As a result, I may experience the thought as having been put into my head by someone else—a delusion of thought insertion. If the thought 'Catch the bus!' occurs as inner speech, Frith claims, I will experience it in the form of a hallucination in which someone is *telling* me to catch the bus.

The examples above capture many of the significant features of Frith's model. The model provides parallel accounts of a number of schizophrenic symptoms, is testable, and has been supported by some recent data. For example, Malenka et al. (1982) and Frith and Done (1989) argue that self-monitoring in the domain of motor-behaviour allows subjects to make rapid corrections of motor errors. Schizophrenic subjects who suffer from delusions such as thought insertion fail to make such corrections, suggesting that their delusions may also reflect a failure of self-monitoring.

4. Egocentricity

Frith's model offers a functional explanation of the development of delusion. What does this account say about the nature of their irrationality? We propose that delusional irrationality consists in a failure of what we will call *egocentricity*, a property of thought closely related to the self-monitoring central to Frith's theory. Self-monitoring is the cognitive process of representing an intention; egocentricity is the property a thought gets in virtue of being so monitored. It is the property of the thought which allows the thinker of that thought to recognize it as having originated in his mind.

To illustrate egocentricity, consider the following case due to John Perry (1979).⁸ Suppose I am pushing my cart down the aisle of a supermarket and notice that someone is producing a trail of sugar, presumably from a leaking bag. I move down the aisle and back up again in search of the person whose bag of sugar is leaking in order to tell him that he is making a mess. The trail gets thicker and thicker, but I am unable to catch up with this person. Eventually it dawns on me that *I* am the person whose bag of sugar is leaking, and, on realizing this, my behavior changes; I stop chasing the unknown person whose bag of sugar is leaking and close the bag of sugar in my own cart.

The central aspect of this case is the transition from my awareness that *someone* has a leaking bag of sugar to the awareness that *I* am that person. When I know that someone has a leaking bag of sugar, there is a sense in which I have a complete description of the situation in the supermarket aisle. What I fail to know is that *I* am the person whose bag of sugar is leaking. What I am missing is information about my role—my conceptual location, so to speak—in the situation.

Something analogous seems to be true of delusions in schizophrenia. If Frith's model is correct, then in thought insertion the subject has the thought 'Catch the bus!', but because there is no monitoring of willed intention, the subject does not have available to him an explanation of the source of that thought—an explanation of the form 'I have the intention to be on time which

⁸ We put Perry's case to a somewhat different use from that of Perry himself, however (see also Evans, 1982).

naturally led to the thought to catch the bus. The intention was mine and so, therefore, is the thought.⁹ While the subject is quite aware that the thought resides in his mind, he is not aware that it originated there. In effect, the subject is capable of thinking ‘Someone has produced the thought “Catch the bus!”’, but not ‘I am that someone’. There is an awareness that a thought has been produced, as well as an awareness of the content of the thought, but no awareness of one’s own role as its producer.¹⁰ As in Perry’s case, the deluded subject has third-person, but not egocentric, information to the effect that the thought of which he is aware originated in his own mind. One way of conceiving of the failure of self-monitoring, therefore, is as a state of affairs in which the subject’s thoughts fail to have the property that allows them to be identified as having originated in his mind. The thinker is thus in no doubt about who he is, or about the thoughts he is having. Rather, he fails to know that he produced those thoughts.¹¹ Because we have claimed that the delusions of schizophrenia are paradigm cases of irrationality and that a failure of egocentricity is the cause of those delusions, we take egocentricity to be a condition on rationality.

5. Does Egocentricity Belong in the Departments of Procedural or Content Rationality?

We turn now to the question of whether a failure of egocentricity belongs to procedural irrationality, content irrationality, or neither. On the face of it, egocentricity seems to bear no relation either to procedural or content conceptions of rationality. Nevertheless, the defender of the traditional departments will argue as follows. Even if one believes that egocentricity plays a causal role in the production of delusion, it does not follow that the violation of egocentricity is constitutive of delusional irrationality. One must distinguish between the causal sources of delusion and its psychological explanation on the one hand, and the conditions which are constitutive of the irrationality on the other. It is quite possible that a unique form of irrationality—supposing that that is what schizophrenic delusion is—may nevertheless belong with other forms of irrationality in one of the traditional departments.

⁹ Notice, however, that questions remain about this aspect of the model. Why is it that an awareness of a thought is insufficient by itself to identify the thought as one’s own but that the presence of an intention is sufficient? What enables the subject to identify the intention as his own? Why, in other words, is there no delusion of intention insertion?

¹⁰ Notice the distinction between two potential mental confusions. I can be aware that someone has produced a thought, but be unaware that I am the person who has produced it. In contrast, I might be aware that John Perry has produced a thought, but be unaware that I am John Perry. The latter case may be relevant to explaining certain delusions of self-misidentification (see Breen et al., this volume).

¹¹ See Peacocke, 1999, who also addresses Frith’s conception of self-monitoring and its implications for self-knowledge: ‘A conception of ownership which involves origination is crucial for rational thought and action’ (p. 244).

We accept the distinction between causal factors in delusion and the correct classification of the irrationality so produced. We must show, therefore, that procedural and content accounts of delusion are not adequate to the phenomena. Since our approach is naturalistic, the correct description of schizophrenic delusion is the one that provides the best characterization of the phenomena overall. We argue that, however plausible procedural and content theories are, they fail to explain all of the features of the delusional state and hence that a different notion of irrationality is necessary to capture those features. Because it is the violation of egocentricity that leads to schizophrenic delusion, we conclude that the study of egocentricity belongs in that department as well.

5.1 Procedural and Content Accounts¹²

The classical accounts of rationality would characterize our central case of thought insertion as a failure of procedural rationality. On these accounts, a violation of a rule or a principle of reasoning is responsible for the deluded state. Content rationality, in contrast, locates delusion in the ineligibility of delusional belief. We consider a variety of such explanations below.

5.1.1 Reflective Equilibrium Thought insertion may be caused by a violation of a principle of belief revision. Whatever the local plausibility of thought insertion as an explanation of a thought whose origin is unknown, it is wildly implausible against the background of our culture's common beliefs. If one assumes that the subject shares many of these beliefs, he violates a principle of *reflective equilibrium* (e.g. Harman, 1986), according to which one should aim for maximal coherence among one's beliefs. Reflective equilibrium would require rejecting thought insertion in favour of an explanation that cohered better with one's set of background beliefs, including those shared by the culture at large. In general, any account of rationality that requires coherence among beliefs will be violated by the hypothesis of thought insertion.¹³

5.1.2 Evidence A closely related possibility is that thought insertion may violate a principle of relevant evidence. Stone and Young (1997; see also Young, this volume) have argued that subjects with Capgras delusion are susceptible to biases in explanation. That is, they are apt to ignore evidence—in Capgras delusion, for example, of how unlikely it is that a loved one has been replaced by a double—and to give undue weight to other kinds of evidence. Since the delusion fails to be supported by legitimate evidence, an evaluation

¹² In psychiatry, psychosis is divided into two disorders: that of thought process and that of thought content (Kaplan and Sadock, 1995). Though related, these categories do not correspond to procedural and content rationality.

¹³ Langdon and Coltheart, this volume, argue that a reasoning deficit in this area is one of the factors involved in producing delusions.

of this evidence should lead the schizophrenic subject to reconsider the delusional belief.¹⁴

5.1.3 Methodology A third procedural violation that might be invoked is a methodological one. The mere presence of a thought the origin of which is inaccessible to the subject is not itself sufficient for the production of a delusion. Suppose one were to think to oneself 'Catch the bus!' but be unable to explain the presence of that thought. One could choose to be puzzled by this occurrence without moving on to providing an explanation of it.¹⁵ Such cases are commonplace. One wonders why a particular thought has popped into one's head; or one starts to daydream about something quite disconnected from the present conditions; or one gets an urge to go out into the sunshine; and so forth. In such cases, one rarely stops to wonder about the origins of these thoughts. Similarly, there are normal confusions, particularly of memory, in which one cannot remember, for example, whether a particular thought or image is a memory of a real event or of a dream. Further, illusions of memory occur just as perceptual illusions do: when an event is vividly imagined, the imagined event is often remembered as having occurred (Goff and Roediger, 1998), but one is not in general puzzled, or even aware, of these occurrences. One could, therefore, explain thought insertion as resulting from the violation of a methodological principle advocating a *suspension* of explanation when no reasonable explanation is available.¹⁶

There are cases more prosaic than delusion in which abstaining from explanation is the rational thing to do. Suppose, for example, that an experiment generates data that are inconsistent with the best relevant scientific theory. There are a number of things a scientist can do: she can reject the theory; doubt the correctness of the data; do further experiments; or do nothing. Suppose further that the scientist has great confidence in both the theory and in the data. In such a case it may be maximally rational to do nothing. Accepting inconsistency or inexplicability may often present the best choice under the circumstances,¹⁷ and adopting a rule of reasoning according to which some

¹⁴ Langdon and Coltheart, this volume, argue that the bias account is inadequate because biases can eventually be overcome. They therefore suggest a deficit account of delusion.

¹⁵ We thus differ from Langdon and Coltheart, this volume, in holding that the constraints of rationality do not *require* certain experiences to be explained.

¹⁶ Notice that one can distinguish two senses of an explanation's being 'available' to a subject. On one sense of 'available', normal subjects in the same circumstances would be expected to have the cognitive ability to produce the relevant explanation. On a second sense, an explanation is available if a subject *ideally placed* would be able to appeal to it. Schizophrenic subjects are abnormal because certain explanations are not available to them in the first sense.

¹⁷ This suggestion was explored by John Bigelow in an Australian Association of Philosophy presidential address: Armidale, Australia, 1995.

explanation must be given in *all* circumstances would no doubt dispose one—and could dispose schizophrenics—to produce irrational explanations.

5.1.4 Content-ineligibility Given the bizarre quality of thought insertion as a belief about one's experience, a content explanation looks promising. On a content account, a delusion will be said to be irrational because, as Lewis (1986) puts it, 'it assigns ineligible, unreasonable content when a more eligible assignment would have fit behavior equally well' (p. 39). Further, the schizophrenic subject manages to incorporate the implausible belief into his system of delusion in a way that is reminiscent of Lewis's description of the 'twisting' of belief and desire, namely, by a *system* of beliefs each of which is highly improbable or impossible on its own but which seems to acquire a veneer of plausibility from its coherence with the others.

Content irrationality may also offer a more parsimonious explanation of delusion than procedural accounts can. It is a familiar experience to reason oneself into a belief that seems plausible in the context of that reasoning but, subsequently, and in the absence of the reasoning context, seems quite implausible. The defender of procedural irrationality has to explain not only how delusion comes about but also why the delusion fails to be rejected upon review. If, however, the irrationality of delusion is located in the tendency of schizophrenic subjects to be insensitive to bizarre contents of thought, then there is no need to posit a procedural deficit in both the reasoning process and the subsequent process of evaluation but only a single deficit of 'bizarre content blindness'.

5.2 Problems with These Accounts

In our view, none of these accounts provides a satisfactory naturalistic explanation of schizophrenic delusion. We consider the difficulties with each account in turn.

5.2.1 Reflective Equilibrium Reflective equilibrium offers an unsatisfactory explanation of the irrationality of thought insertion because if the schizophrenic agent has a global impairment in his ability to maintain coherence among his beliefs, one would expect this impairment to be domain-general. An inability to achieve reflective equilibrium among beliefs would infect the agent's whole system of beliefs, and one would expect the agent to experience delusions of every imaginable kind. But the delusions of schizophrenic subjects tend to be restricted to the social domain (Bentall et al., 1991; Frith, 1992).¹⁸

¹⁸ One possible response to this argument, therefore, is that delusions all fall into a single category of *social reasoning* and, therefore, that the boundaries within which the errors of reasoning fall are natural ones (see Bentall et al., 1991). This claim could be developed in the context of a defence of the existence of a module for 'mentalizing', that is, for representing the mental states of others (see Currie, this volume).

Even if the proceduralist were to hypothesize that the schizophrenic has an impairment of the ability to apply reflective equilibrium only in this restricted domain, delusional states tend to revolve around highly specific and predictable ideas and are therefore much rarer than delusion ought to be if schizophrenics have a deficiency of reflective equilibrium in the domain of the social. The classic Schneiderian taxonomy (Schneider, 1959) includes seven types of delusion as diagnostic of schizophrenia; a more recent classification identifies only five (Cutting, 1995). Given this small number of species of delusion, it is hard to see how one could identify a natural cognitive domain, social or otherwise, that would isolate all and only those topics that emerge in schizophrenic delusion, and it would be quite *ad hoc* to posit a defect of reflective equilibrium that applies only to those topics by definition. No such hypothesis is likely to be supported by a scientific theory of schizophrenic pathology.

5.2.2 Evidence A principle of evidence evaluation suffers from the same problem as that of reflective equilibrium. A global problem within the web of belief should lead to innumerable false beliefs, delusions, and, indeed, to behavioural chaos, but this does not occur in schizophrenia. Schizophrenic subjects appear to be just as capable of using the presence of a sign saying 'Metro' as evidence that a metro is nearby. In addition, a hallmark of the ability to evaluate evidence is the capacity to *change* one's view based on the introduction of new evidence (see Langdon and Coltheart, this volume). This too is something that schizophrenic subjects do not seem to have trouble doing: new evidence to the effect that the Angrignon metro goes west and not east will lead the eastward-travelling schizophrenic to change lines. In contrast, delusional states are, by definition, highly stable and resistant to revision. Given that schizophrenics do not have a general difficulty with evidence, it is unlikely that this stability lies in an insusceptibility to evidence. And, again, the hypothesis of a *local* failure to be responsive to evidence suffers from the problem of identifying a cognitive domain that captures all and only the delusions of interest. If the hypothesized domain is too big, the hypothesis will fail to explain the specificity of delusion; if it is too small, it will fail to carve the cognitive domain at its joints.

5.2.3 Methodology This version of the procedural account is compelling because it appears at first blush to be based on a minimal norm of rationality and one that is very easy to obey: when no plausible explanations are available, suspend belief. But, in fact, this requirement on rationality is much more demanding than it first appears. We claimed that when a scientist is faced with data that conflict with theory, it may be maximally rational to do nothing. Is this a sensible course of action for the schizophrenic? It is natural to think of the schizophrenic subject's choice as one between believing that the CIA is inserting thoughts into his head and believing nothing, and the option of believing nothing seems far more appealing to the outside observer. But this

is not a correct description of the options. The schizophrenic in fact faces the choice of having some account of what is happening to him—some account of the structure of his experience—and being utterly at a loss to understand that experience. In asking the schizophrenic subject to refrain from explaining his experience, we would be asking him to forgo any story about why his experience is so strange, and that seems to be a demand that no agent—at least no agent that approximates actual human agents—could meet. Such a demand would be little different from asking Joseph K., the central character of Kafka's (1925/1998) *Trial*, to stop wondering about why he's been arrested. Given that our account is naturalistic, it should not posit a violation of a norm that no actual agent could meet.

Even in the scientific domain, it is not clear that such a constraint is plausible. The scientist in our case above is faced with a conflict between the data of an experiment and the theory that is relevant to that data. However, the scientific analogy to schizophrenia would rather be to giving up any account at all of a large and significant aspect of the world. Wilfrid Sellars (e.g. Sellars, 1963) conceived of science as a way of locating oneself in the world, and this requires having at least general working hypotheses. In addition, such hypotheses are also necessary to identify new lines of research. Thus, the schizophrenic subject might actually be *violating* a norm of scientific method by refraining from a working hypothesis with which to understand their experience.

There is a further virtue of adopting a delusional account of one's experience rather than no account at all. One assumes that delusional experience—at least in the case of thought insertion, paranoid delusions, and the like—is very unpleasant. A delusional account offers the subject the possibility not only of explaining his experience but of doing something about it, and that sort of motive is a powerful one for normal agents. It seems wrong-headed to claim that a widely shared motive among normal agents constitutes a form of irrationality in schizophrenics. Thus, faced with the strange experiences of schizophrenia, the hypothesis of thought insertion is perhaps one of the more *plausible* suppositions one might make.

Are there other grounds on which the schizophrenic is criticizable? If a suspension of explanation is too demanding a requirement on rationality, could we not demand of the schizophrenic subject that he opt for the hypothesis that he is mentally or neurologically ill? This may still be asking too much. The experience of non-egocentric thought may be so pervasive and compelling that any hypothesis that calls that experience into doubt is automatically ruled out of court. Again, it may be easier, and indeed more sensible, to adopt thought insertion as a hypothesis than to adopt the view that one's experience is systematically incorrect—that one is living in a sort of virtual reality. In addition, to posit mental illness in the face of the systematic strangeness of experience is to give up one's identity as an agent capable of understanding and interacting adaptively with the outside world. This is surely a core background assumption

of the explanation of one's own behaviour and that of others. Thought insertion may seem a less radical option, therefore, than giving up that assumption.

5.2.4 Content-ineligibility It is quite natural to characterize delusion as a pathology of belief and to assume that this pathology is responsible for the irrationality. However, the attempt to reduce the irrationality of delusion to the bizarre contents of delusional beliefs suffers from the same problem as the procedural accounts just evaluated: a content account does not explain why the bizarre beliefs of these subjects seem so narrowly restricted and predictable in content. If schizophrenic delusion were an instance of content irrationality, one would expect that bizarre beliefs would be shot through the system of belief, but this is not in fact the case. In general, schizophrenic subjects believe in eating food in order to satisfy hunger; in going inside to get warm; in taking the metro to get from place to place; and so on.

Further, even if we restrict our attention to the strange experience schizophrenics have and are, presumably, trying to explain, it remains the case that there are, in principle, an indefinite number of beliefs that could be generated on the basis of that experience, many of them less implausible than thought insertion. It is very surprising, therefore, that delusional subjects choose thought insertion as an explanation sufficiently frequently for it to be a hallmark of schizophrenia.¹⁹

There is a further difficulty with the content account. Not all delusional beliefs and desires will turn out to be beliefs or desires that are ineligible in Lewis's sense. What does 'ineligible' mean here? Lewis chooses as paradigms of ineligible contents the desire for a saucer of mud and the belief that emeralds are *grue*—green if examined before some time, *t*, in the future and blue otherwise. What unites both of these mental states is not that they are bizarre, but rather that they fall outside the network of mental states we take to be relevant to the description of the physical world and of human behaviour. In response to someone's expression of a bizarre desire—a desire to sing Pagliacci from the bottom of a herring barrel, for example—one might be inclined to question or dissuade. In response to the expression of a desire for a saucer of mud, one can only be incredulous. Where does one get a purchase on such a desire? Because it falls outside of the framework of human desires, as we understand it, one cannot. The same holds for the belief that emeralds are *grue*. One can argue with someone who believes that emeralds are pink, but one cannot negotiate with the *grue*-ascriber; emeralds *are* *grue*, after all. But *grue* is not a predicate that forms part of the apparatus we use to describe the physical world.

Are delusions ineligible in Lewis's sense? Clearly not. It is bizarre to think that the CIA is tapping your phone. In the light of the relevant evidence, this

¹⁹ Gerrans (this volume) assesses the extent to which Capgras delusion is related to, in our terms, procedural rationality. He also points to the confined nature of these delusions.

belief may be highly improbable, but it is not ineligible. The CIA may *indeed* be tapping your phone; just because you're paranoid, as they say, doesn't mean you don't have enemies. The same applies to our central case of thought insertion. The idea that the FBI has sewn a radio transmitter into your skull and is thereby inserting thoughts into your head is bizarre in the extreme, but it is not outside the bounds of explanations we are willing to consider. Even if the delusion is that thoughts are being inserted by telepathy or through some mysterious ether, we are willing to entertain it; science fiction is filled with scenarios of just this kind.

We conclude, therefore, that plausible procedural and content accounts of delusion are not successful explanations of the nature of the irrationality in schizophrenic delusion. A failure of egocentricity is not therefore either a procedural or a content failure. A different account of the location of the irrationality is required. We claim that the domain-specificity of schizophrenic delusion derives from the fact that the experience of schizophrenics has a content that makes their beliefs quite natural. Our proposal is that it is the *experience* of schizophrenic subjects that is defective, and, for this reason, delusions, and thus irrationality, are an inevitable outcome even if the reasoning procedures applied to those experiences are entirely normal. We will therefore suggest that a new kind of irrationality—experiential irrationality—must be posited to explain delusion in schizophrenia.

6. A New Department of Rationality: Experiential Rationality

Both procedural and content accounts of the irrationality of delusion assume that schizophrenics have strange experiences—on Frith's account, the experience of a thought whose origin is unknown—that are then explained by means of defective procedures of reasoning or with beliefs with ineligible contents (see Breen et al., Langdon and Coltheart, and Young, all in this volume). As we have seen, these accounts locate the irrationality of delusion in the defective reasoning procedures or the bizarre contents of belief. In our view, this is the wrong way to describe the phenomenology of delusion. Unexplained phenomena—such as the occurrence of a stray thought whose origin is unknown—are all around us. If schizophrenic subjects suffered from procedural or content irrationality, one would expect their mental life to be replete with delusions of every imaginable sort. But, as we have argued, this is not the case.

We suggested above that the central feature of the phenomenology of delusion is the alien nature of the subject's thoughts. This is entirely in keeping with Frith's account, according to which schizophrenic thoughts violate egocentricity and are experienced as originating outside of one's mind. We claim, therefore, that the source of thought insertion and related delusions is the experience itself of the schizophrenic subject, and, in particular, its alien quality. The elaboration of the delusion in hypotheses and ancillary beliefs should

be understood to be derivative from, or secondary to, this experience. Thus the violation of egocentricity does not merely produce strange experiences that form the basis of delusional beliefs as the result of pathological processes of thought or reasoning. Rather, it is *the experience of non-egocentric thought as alien* that is the delusion itself. The alien quality of the delusional experience is part of its content, and it is the content of experience that is the locus of the delusion and thus of the irrationality. At least some delusions, therefore, are best explained as *disorders of experience* rather than disorders of belief, desire, or reasoning. Thus whereas Stone and Young (1997) and Langdon and Coltheart (this volume) locate the source of the delusion in the reasoning process that engages with the strange experiences of schizophrenic subjects, we claim that the source of the delusion is the strange experience itself. The verbal reports of patients express the *way* in which the experience is strange.²⁰

Consider thought insertion again as an illustration. A thought appears in the mind of the subject. Because it is not tagged as originating from the subject's own mind, it is experienced as alien—an experience as of something alien having been introduced into one's consciousness. Recall the description given by Kraepelin (1919): “My senses don’t belong to me any more, they are being unlawfully taken from me.” . . . “They take my thoughts from me and nothing comes back but a ragamuffin.” . . . his thoughts are “plundered, organised and published.” “The voices and my brain are one, I must think what the voice says”.’ The way in which a delusion—though not, in this case, the delusion of thought insertion—can be embedded in experience is brought out clearly in this quote from a schizophrenic subject (quoted in Torrey, 1995):

At this point, I panicked and tried to run away, but the attendant in the parking lot seemed to be making a sign to motion me back. I thought I caught brief glimpses of a friend and my wife so I decided to go back into the hospital. A custodian’s eyes attracted my attention; they were especially large and piercing. He looked very powerful. He had to be ‘in on it,’ maybe he was giving medicine in some way. Then I began to have the feeling that other people were watching me. And, as periodically happened throughout the early stages, I said to myself that the whole thing was absurd, but when I looked again the people really were watching me (p. 53).

It seems clear that the experiences of this subject do not lead him to a deluded

²⁰ Note that nothing we say is meant to suggest that the *only* kind of irrationality that presents itself in schizophrenic delusions of self-monitoring is the experiential irrationality brought about by a failure of egocentricity. For example, the fact that schizophrenic subjects do not subsequently revise their views about their delusions in the light of further evidence might require some reference to a form of procedural or content irrationality. Our claim is only that the best explanation of the original delusion itself is the irrationality of experience.

hypothesis but rather that his experience misrepresents the world in a delusional way. Once we suppose that the origin of thought insertion is an experience of certain mental states as alien, the delusions of thought insertion, of control, and of persecution become much less surprising. The experience of an alien presence in one's consciousness is naturally expressed in these delusional states and in the system of beliefs that is built upon it. Once egocentricity is violated, the particular irrationality of these delusions is almost inevitable.

As we suggested above, in the light of schizophrenic experience, delusion can be seen as a sensible cognitive response. To refrain from explanation is, we have suggested, psychologically impossible, and the other options, such as hypothesizing that one is mentally ill, may require giving up beliefs that are more important to the agent—such as the belief that he *is* an agent; that he has some grip on the nature of his experience; or that he can affect his environment. Thought insertion, however bizarre, may be among the more adaptive hypotheses one could adopt. This seems particularly true if one supposes that delusional experience is pervasive in schizophrenia and that more 'rational' hypotheses than thought insertion would require dismissing the veridicality and relevance of one's experience. Certainly, normal subjects have a strong inclination to take most of their experience at face value, and one would expect the same of schizophrenic subjects. The subject quoted above, for example, resists the idea that people are watching him, but then his experience seems to overwhelm him: 'I said to myself that the whole thing was absurd, but when I looked again the people really were watching me.'

Further support for the experiential view derives from its continuity with the fact of perceptual abnormality that characterizes almost two-thirds of schizophrenic subjects (Torrey, 1995). Indeed, perceptual abnormality—such as heightened sensitivity to stimuli—is the most invariant feature of the early stages of schizophrenia (Cutting and Dunne, 1989). Given the pervasiveness of perceptual dysfunction, a theory of delusion that focuses on experience is more parsimonious than procedural or content accounts. It posits an abnormality in experience and is thus in the spirit of the proposal by Maher (1974), who locates the source of the delusion in abnormal perception.

The parsimony of the experiential view is enhanced when schizophrenic hallucination is considered because hallucinations are precisely perceptual experiences that are often intrinsically alien—such as when they are experiences of *other people's voices* commenting, giving a command, or whatever. A hallucination is an unusual form of experience, but no subsequent judgement is required on the part of the subject for the experience to become delusional. It has a delusional content, but that content is embedded in an experience rather than in a belief or desire. And, indeed, there is some evidence that hallucinations arise as a result of a violation of egocentricity in the domain of inner speech (Stein and Richardson, 1999), thus raising the possibility of a

unification of delusion and hallucination in schizophrenia by means of the concept of egocentricity.

Finally, this account of delusion meshes well with a related symptom of schizophrenia, namely, the fact that schizophrenics sometimes experience their bodies, or body parts, as alien. One patient (quoted in Torrey, 1995) describes the experience this way:

I get shaky in the knees and my chest is like a mountain in front of me, and my body actions are different. The arms and legs are apart and away from me and they go on their own. That's when I feel I am the other person and copy their movements, or else stop and stand like a statue. I have to stop to find out whether my hand is in my pocket or not. I'm frightened to move or turn my head. Sometimes I let my arms roll to see whether they will land (p. 65).

The experience of body alienness also occurs in *anosognosia*, a lack of awareness of a sensory or neurological deficit (see Goetz and Pappert, 1999).²¹ An extreme version of this condition is the *alien hand syndrome*. Kaufman (1995) describes the condition as follows:

In this disorder . . . a patient's left²² hand retains some rudimentary motor and sensory functions, but they cannot be appreciated by the rest of the patient's brain. Without the patient's awareness, the hand moves semipurposefully, makes its own explorations, and performs simple tasks, such as scratching and moving bedclothes. In a unique, often quoted example, a patient reported that her hand was attempting to choke her. The alien hand syndrome rests on the patient having at least two misperceptions: (1) the patient does not possess the hand and (2) the hand's movements are independent or governed by another person (the alien). Most patients feel divorced from the hand or, at most, express a tenuous attachment to it (pp. 182–3).

²¹ Intriguingly, patients with damage to the left motor cortex are able to explain an inability to move a limb by appealing to their brain damage, but patients with right-cortical damage claim that they are unable to move a limb because it belongs to someone else. On our account, a possible explanation is the following. In left-hemisphere damage, motor commands are given to move the limb and are monitored. When the limb fails to move, the experience is of an inability to execute this command. In right-hemisphere damage, the motor command is not tagged as originating from the subject and so the experience of the limb is as of its belonging to someone else. After all, in normal circumstances a limb to which one cannot give commands *should* be experienced as belonging to someone else. (We are grateful to Max Coltheart for alerting us to this issue.)

²² The condition typically occurs in individuals with brain damage to a non-dominant right hemisphere. The incapacity is therefore to the left hand. (See note 21.)

The analogy with schizophrenic delusion is striking. In this case, however, it seems especially plausible that the idea that the hand is controlled by an alien stems from the subject's experience of the hand as disconnected from the subject's consciousness, both in the sensory and the motor domain, rather than from a belief or hypothesis formed on the basis of the fact that the hand is not under the subject's control. It is thus not surprising that a thought not identified as originating in one's own mind—like a limb not identified by sensation as belonging to one's own body—*feels* alien quite in the absence of any reasoning, or hypothesis-formation, about that thought.

A failure of egocentricity, therefore, leads to a mental state that is irrational even though there may be no irrational belief or desire content nor any explicit hypothesis (whether conscious or unconscious) that is invoked to explain the experience. Although the form of irrationality produced by a violation of egocentricity is not reducible to procedural or content irrationality, it is also a worthy deserver of the name 'irrational'. Because these irrational states do not fall within the traditional categories, we propose a new department of rationality called *experiential rationality*.

We note, finally, that if the present analysis is correct and there are experiential forms of irrationality, we expect that further investigation into the phenomena of psychopathology will reveal other constraints on rationality that do not fall within the departments of procedural or content irrationality.

*Institute of Advanced Studies
Australian National University*

References

- American Psychiatric Association. 1994: *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV). Washington, DC: American Psychiatric Association.
- Anscombe, G.E.M. 1957: *Intention*. Oxford: Blackwell.
- Bentall, R.P., Kaney, S. and Dewey, M.E. 1991: Paranoia and social reasoning: An attribution theory analysis. *British Journal of Child Psychology*, 30, 13–23.
- Brown, H.I. 1988: *Rationality*. London: Routledge.
- Cherniak, C. 1986: *Minimal Rationality*. Cambridge, MA: MIT Press.
- Cutting, J. 1990: *The Right Cerebral Hemisphere and Psychiatric Disorders*. Oxford University Press.
- Cutting, J. 1995: Descriptive psychopathology. In S.R. Hirsch and D.R. Weinberger (eds), *Schizophrenia*. Oxford: Blackwell Scientific.
- Cutting, J. and Dunne, F. 1989: Subjective experience of schizophrenia. *Schizophrenia Bulletin*, 15, 217–31.

- Evans, G. 1982: *The Varieties of Reference*. Oxford University Press.
- Feinberg, I. 1978: Efferece copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin*, 4, 636–40.
- Feinberg, I. and Guazzelli, M. 1999: Schizophrenia: A disorder of the corollary discharge systems that integrate the motor systems of thought with the sensory systems of consciousness. *British Journal of Psychiatry*, 174, 196–204.
- 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.
- Frith, C.D. and Done, D.J. 1989: Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychological Medicine*, 19, 359–63.
- Garety, P.A., and Freeman, D. 1999: Cognitive approaches to delusions: A critical review of theories and evidence. *British Journal of Clinical Psychology*, 38, 113–54.
- Goetz, C.G. and Pappert, E.J. 1999: *Textbook of Clinical Neurology*. Philadelphia: W.B. Saunders.
- Goff, L.M. and Roediger, H.L. 1998: Imagination inflation for actions and events: Repeated imaginings lead to illusory recollections. *Memory and Cognition*, 26, 20–33.
- Goodman, N. 1973: *Fact, Fiction, and Forecast*, 3rd edn. Indianapolis: Bobbs-Merrill.
- Harman, G. 1986: *Change in View*. Cambridge, MA: MIT Press.
- Hume, D. 1888/1978: *A Treatise of Human Nature*, 2nd edn, ed. L.A. Selby-Bigge and P.H. Nidditch. Oxford University Press.
- Kafka, F. 1925/1998: *The Trial*, trans. B. Mitchell. New York: Schocken Books.
- Kahneman D., Slovic P. and Tversky, A. 1982: *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge University Press.
- Kaplan, M. 1996: *Decision Theory as Philosophy*. Cambridge University Press.
- Kaplan, H.I. and Sadock, B.J. (eds). 1995: *Comprehensive Textbook of Psychiatry*, 6th edn. Baltimore: Williams and Wilkins.
- Kaufman, D.M. 1995: *Clinical Neurology for Psychiatrists*, 4th edn. Philadelphia: W.B. Saunders Company.
- Kraepelin, E. 1919: *Dementia Praecox*, trans. R.M. Barclay, ed. G.M. Robertson. Edinburgh: E. & S. Livingstone.
- Lewis, D. 1986: *On the Plurality of Worlds*. Oxford University Press.
- Maher, B. 1974: Delusional thinking and perceptual disorder. *Journal of Individual Psychology*, 30, 98–113.
- Malenka, R.C., Angel, R.W., Hampton, B. and Berger, P.A. 1982: Impaired central error correcting behaviour in schizophrenia. *Archives of General Psychiatry*, 39, 101–7.
- Peacocke, C. 1999: *Being Known*. Oxford University Press.
- Perry, J. 1979: The problem of the essential indexical. *Noûs*, 13, 3–21.
- Rescher, N. 1988: *Rationality*. Oxford University Press.
- Ryle, G. 1949: *The Concept of Mind*. London: Hutchison & Company.

- Schneider, K. 1959: *Clinical Psychopathology*, trans. M.W. Hamilton. New York: Grune & Stratton.
- Sellars, W. 1963: Science and the manifest image of man. In W. Sellars, *Science, Perception, and Reality*. New York: Routledge and Kegan Paul.
- Skyrms, B. 1986: *Choice and Chance*, 3rd edn. Belmont: Wadsworth.
- Stein, J. and Richardson, A. 1999: Cognitive disorders: A question of misattribution. *Current Biology*, 9(10), R374–6.
- Stone, T. and Young, A.W. 1997: Delusions and brain injury: The philosophy and psychology of belief. *Mind and Language*, 12, 327–64.
- Torrey, E.F. 1995: *Surviving Schizophrenia*, 3rd edn. New York: Harper Perennial Press.