British Journal of Clinical Psychology (1999), **38**, 113–154 *Printed in Great Britain*

Cognitive approaches to delusions: A critical review of theories and evidence

Philippa A. Garety*

Department of Clinical Psychology, GKT, University of London, St Thomas' Hospital, UK

Daniel Freeman

Department of Clinical Psychology, Institute of Psychiatry, University of London, UK

Purpose. To review critically the evidence for three contemporary theories of delusions.

Methods. The theoretical approaches to delusions proposed by Frith and colleagues ('theory of mind' deficits), Garety and colleagues (multi-factorial, but involving probabilistic reasoning biases) and Bentall and colleagues (attributional style and self-discrepancies) are summarised. The findings of empirical papers directly relevant to these proposals are critically reviewed. These papers were identified by computerised literature searches (for the years 1987–1997) and a hand search.

Results. The evidence does not unequivocally support any of the approaches as proposed. However, strong evidence is found to support modifications of Garety and colleagues' and Bentall and colleagues' theories. Studies have replicated a 'jumping to conclusions' data-gathering bias and an externalising attributional bias in people with delusions. There is preliminary evidence for a 'theory of mind' deficit, as proposed by Frith, although possibly related to a more general reasoning bias. Evidence for an underlying discrepancy between ideal and actual self-representations is weaker.

Conclusions. A multi-factorial model of delusion formation and maintenance incorporating a data-gathering bias and attributional style, together with other factors (e.g. perceptual processing, meta-representation) is consistent with the current evidence. It is recommended that these findings be incorporated into cognitive therapy approaches. However, there are limitations to existing research. Future studies should incorporate longitudinal designs and first episode studies, and should not neglect the co-morbidity of delusions, including affective processes, or the multi-dimensional nature of delusions.

'Controversy and disagreement are always at the heart of scientific advance. The enterprise is driven by Karl Popper's mixture of bold conjectures followed by rigorous attempts at falsification; competition in the world of ideas may be the best (perhaps the only) example of how

* Requests for reprints should be addressed to Professor P. A. Garety, Department of Clinical Psychology, Guy's, King's and St Thomas' School of Medicine, University of London, St Thomas' Hospital, Lambeth Palace Road, London SE1 7EH, UK.

> a free market can produce (not the perfect theory but at least) successive approximations to the Truth. Discord, then, is healthy.' *Marshall & Halligan (1996, p. 3)*

These are the opening lines of a recent book about mental illness or 'madness', which consists of case studies of a variety of delusions and hallucinations. Marshall & Halligan recognise how wide a range of opinion there exists in the contemporary literature over such fundamental questions as the value of a diagnostic approach to 'schizophrenia' and what kinds of theoretical interpretation of positive and negative symptoms are possible and plausible. They commend debate and disagreement as central to the development of understanding. They also recommend the individual symptom approach, in contrast to a diagnostic one, as offering greater promise for the future; it is just this approach which has come to prominence over the past decade in psychological approaches to delusions.

In this study the authors propose to review and offer an interim appraisal of this recent work, which has used a symptom approach to investigate delusions, considering both the theories proposed and the evidence so far amassed. They take as their focus delusions in general, although some studies consider persecutory delusions in particular. They restrict the review to cognitive theories and empirical studies of delusions, where the underlying construct is a model of performance found in the general population in the cognitive domain of interest (i.e. models of belief formation and maintenance). Their main question is: which cognitive processes, that have been investigated, are associated with the occurrence of delusions? Ultimately, at least one goal of such cognitive investigations is to specify what leads to the development and maintenance of delusions in order to inform therapeutic change, possibly, but not necessarily or exclusively, by means of cognitive therapy. The growing evidence of the effectiveness of cognitive-behavioural approaches with delusions adds a further impetus to the task of theoretical development (Bouchard, Vallières, Roy & Maziade, 1996; Kuipers *et al*., 1997).

In recent years there have been three main theoretical approaches which have stimulated empirical studies: the work of Frith, Garety and Bentall and their respective colleagues. The present authors consider each of these accounts in detail and evaluate the empirical support for each theory, before addressing a number of other issues. These include the extent to which the cause and maintenance of delusions are addressed, the theories' specificity, whether the three approaches are compatible, contradictory or complementary, and their implications for therapy. First the concept of delusion is introduced and an earlier account of delusions proposed by Maher (1974, 1988) is reviewed.

Delusions

Delusions have traditionally been regarded as fixed, false beliefs, held with absolute conviction and not amenable to reason (e.g. DSM IIIR, American Psychiatric Association (APA), 1987; Mullen, 1979). However, research has promoted a number of changes in the conceptualization of delusions. Empirical investigations have demonstrated that delusions are dimensional (for example, levels of conviction are not always absolute and may fluctuate), multi-dimensional (so that recovery can be determined by change in one of a number of different dimensions—conviction, preoccupation, distress and action) and are not always fixed and unresponsive to

evidence or reason (so that some people with delusions are able to reason about them and can be responsive to contradictory ideas or experiences) (Brett-Jones, Garety & Hemsley, 1987; Chadwick & Lowe, 1990; Garety & Hemsley, 1987; Oltmanns & Maher, 1988; Strauss, 1969). The revised DSM IV (APA, 1994) definition reflects at least one of these changes in acknowledging that delusions may show varying levels of conviction.

Although there are widely acknowledged problems with standard psychiatric definitions, most empirical studies have used operational criteria for the presence of delusions based on them. In practice, a person is found to hold a delusion if the belief is asserted with high conviction; if it appears unlikely to respond (rapidly and significantly) to evidence or reason; and if it fits one of the content categories to be found in diagnostic manuals (such as the Present State Examination (Wing, Cooper & Sartorius, 1974) or its more recent update, the SCAN (World Health Organisation, 1992)). The presence of distress, preoccupation or action accompanying the belief also increases the probability that it will be described as delusional (Oltmanns, 1988). It is therefore quite possible, in the absence of tightly defined operational criteria, that there may be important differences in the characteristics of the 'delusions' reported in the literature which are reviewed below. Although the single symptom approach is thought to bring the advantage of the avoidance of the misclassification of subjects (Persons, 1986), it can be seen that this does not necessarily follow. To what extent this issue poses a problem for this research is an empirical question.

Delusions as explanations of experience

Maher (1974) offered a cognitive account of delusions which emphasised disturbances of perception. He proposed that a delusional individual suffers from primary perceptual anomalies, fundamentally biological in nature, that involve vivid and intense sensory input. This may be an experience of increased vividness of colours, or a difficulty in attending selectively to an auditory stimulus against background noise because of increased prominence of the latter. Some of these anomalies are experienced as hallucinations. Maher argues that the individual, being prone to experience these abnormal percepts, seeks an explanation which is then developed through normal cognitive mechanisms. The explanation (i.e. the delusion) is derived by processes of reasoning that are entirely normal. He argues that the delusion is maintained in the same way as any other strong belief: just as scientists are resistant to disconfirmation of their theories, so are deluded people equally resistant. Furthermore, he suggests that delusional beliefs are reinforced by the anxiety reduction which accompanies the development of an explanation for disturbing or puzzling experiences.

Maher (1988) cites evidence in support of this view from two sources. First, it is noted that delusions occur in a large number of medical and psychological conditions which he argues indicates that delusions serve an adaptive function, secondary to whatever disturbance is caused by the pathogenic agent (Maher & Ross, 1984). Secondly, evidence is cited that irrational beliefs can be provoked in the general population under anomalous environmental conditions, for example undetected hearing loss leading to paranoid ideas (Zimbardo, Andersen & Kabat, 1981). As a

third point, Maher argues that there is an *absence* of evidence for any impairment of reasoning ability, 'apart from the inference made from the presence of the delusions themselves' (Maher, 1988, p. 23).

Maher's account is an elegant single factor model of delusions, in which an abnormality of perceptual processing combined with paradoxically normal reasoning leads to the delusion. It is plausible where delusional beliefs accompany a readily detectable unusual sensory experience, particularly with a known pathology. However, it does not provide a complete account of all delusions. First, some delusions are found to occur in the absence of any anomalous experiences (e.g. Chapman & Chapman, 1988). Secondly, there is growing evidence of reasoning and attributional biases in people with delusions which suggests they may display systematic differences in cognitive processes from those in the general population. Thirdly, it is probable that the experience of anomalous percepts (e.g. hallucinations) is a less passive process than Maher suggests, and may result, in part, from biased cognitive processes in the task of 'reality discrimination' (Slade & Bentall, 1988). However, even if Maher's account applies only to a restricted set of delusions, his proposal is important in positing that delusions are *explanations of experience*, that they represent the individual's attempt *to make sense* of events. This contrasts very markedly with the conventional wisdom regarding delusions, dating from Jaspers (1913), that (primary) delusions are 'ununderstandable' and psychologically irreducible. Arguably, although there exists a scattering of interesting writings on the cognitive processes involved in delusions throughout the twentieth century (see Arthur, 1964; Garety & Hemsley, 1994; and Winters & Neale, 1983 for reviews), Maher's work was the key stimulus for the renewed interest in generating psychological accounts of delusions since the 1980s.

DEFICITS IN META-REPRESENTATION (FRITH)

Frith (1992) has proposed that delusions of reference and of persecution (and third person auditory hallucinations) arise from an inability to represent the beliefs, thoughts and intentions of other people, a 'theory of mind' deficit. The argument can be put concisely: 'By their very nature, delusions of reference, misidentification and persecution are characterised by a misinterpretation of another person's behaviour or intentions. Thus the argument that these symptoms arise as a result of a deficit in a system which enables us to infer what is in the minds of other people is straightforward' (Corcoran, Mercer & Frith, 1995, p. 6). Frith and colleagues also propose that the asocial behaviour and blunted affect seen in patients with negative features reflects an impairment in 'theory of mind', which is thought to predate the onset of the disorder and may reflect an early developmental disruption.

This theory developed from Frith's (1987) earlier suggestion that a failure of the system which monitors one's own actions and their preceding intentions (i.e. a *self-monitoring* deficit) is responsible for the anomalous experiences of thought insertion and of alien control. Thus the experience of thoughts being initiated without any awareness of the intention to initiate them would be described by a person as thought insertion. Similarly actions would appear to be determined by external forces if there was no awareness of the intention to act. Support for this deficit in self-monitoring

comes from an experimental study in which patients with experiences of alien control were less likely than controls to correct errors they had made, indirectly suggesting that they had impairments in monitoring their own actions (Frith & Done, 1989). Since this aspect of Frith's theory primarily concerns anomalies of experience rather than delusional beliefs, this is not considered further.

Frith's model makes clear predictions about the contrasting performance of patients with schizophrenia on tasks which require theory of mind skills. These predictions are stated by Corcoran *et al*. (1995). They predict poor performance from patients with negative symptoms, patients with incoherent speech and patients with delusions of reference and persecution. Patients with passivity experiences are thought to have problems with their own mental states rather than those of others and are predicted to perform such tasks normally. Patients in remission are predicted to perform normally.

Theory of mind studies

In seven recent papers 'theory of mind' studies are reported (see Table 1): four are by Frith and co-workers and three are by other research groups. In Frith and colleagues' first three studies, the same basic design was employed. Patients with a diagnosis of schizophrenia were hierarchically classified by current symptoms into subgroups, from negative symptoms at the top of the hierarchy to 'in remission' at the bottom (see Table 1). The presence of any symptom from a group higher in the hierarchy automatically placed the patient into the higher group. This means that the groups also probably represent increasing levels of symptomatology and that the higher symptom groups are more mixed. (In two studies (Corcoran, Cahill & Frith, 1997; Frith & Corcoran, 1996), the top two groups, negative symptoms and incoherence, are combined into one group, 'behavioural signs', and Corcoran *et al*. (1997) combined the lowest two groups, atypical delusions and in remission.) In all three studies, there were two control groups: a 'normal' control group and a smaller, mixed psychiatric control group. Measures of IQ were taken. The studies employed a variety of different tasks, both verbal and non-verbal (visual), all requiring the participants to make inferences about the mental states of others.

Both 'first order' and 'second order' theory of mind tasks are used. Successful completion of first order tasks requires appreciation of a character's beliefs about the world, while second order tasks involve more complex theory of mind skills and require inferences about a character's beliefs about the beliefs of another character. An example of a first order task item, taken from the 'Hinting Task' devised by Corcoran *et al*. (1995), requires participants to infer the real intentions behind indirect speech utterances:

> 'Paul has to go to an interview and he's running late. While he's cleaning his shoes he says to his wife, Jane: 'I want to wear that blue shirt, but it's very creased.'
>
> Question: What does Paul really mean when he says this?
>
> Extra information: Paul goes on to say 'It's in the ironing basket.'
>
> Question: What does Paul want Jane to do?

Table 1. Theory of mind studies

Study	Task	Groups of participants (*N*)	Results: T.O.M. deficits by group	Results: T.O.M. deficit SCHIZ	Results: T.O.M. deficit NEG	Results: T.O.M. deficit PARA	Comments
1. Corcoran *et al.*, 1995	'Hinting' (infer intentions behind indirect speech)	1. Schizophrenia (55) in subgroups:	√	√	√	√	The combined schizophrenia group perform worse than control groups. The groups were not matched for IQ and IQ is correlated with performance on task in schizophrenia group, but not in controls. The negative symptom subgroup shows clearest deficit and paranoid symptom group shows a deficit relative to controls.
		(i) Negative symptoms (10)	√				
		(ii) Incoherence (3)	√?				
		(iii) Paranoid symptoms (23)	√				
		(iv) Passivity experiences (7)	×				
		(v) Other symptoms (4)	×				
		(vi) In remission (8)	×				
		2. Mixed clinical (anxiety or depression) (14)	×				
		3. Non-clinical (30)	×				
2. Frith & Corcoran, 1996	False belief and deception stories with: 1. Memory test 2. T.O.M. test	1. Schizophrenia (55) in subgroups:	√?	√?	√	√	There are some difficulties interpreting the results because memory or comprehension difficulties and IQ differences have influenced the results. Both behavioural signs group and paranoid group show poorer performance on memory questions. The mixed findings, when groups were matched for IQ, suggest that IQ may make a contribution to T.O.M. deficits. Passivity experience group and patients in remission did not show T.O.M. deficits.
		(i) Behavioural signs of negative symptoms and incoherence (12)	√				
		(ii) Paranoid symptoms (24)	√				
		(iii) Passivity experiences (10)	×				
		(iv) In remission (9)	×				
		2. Mixed clinical (anxiety or depression) (7)	×				
		3. Non-clinical (22)	×				
3. Corcoran *et al.*, 1997	Jokes: 1. Slapstick/physical 2. False belief or deception	1. Schizophrenia (44) in subgroups:	√	√	√?	√?	When IQ was partialled out, the behavioural signs group showed the clearest deficit on T.O.M. but it was also present in the paranoid symptoms group and the passivity experiences group relative to nonclinical but *not* to clinical controls. The clinical controls (mostly depressed) showed a tendency towards a T.O.M. deficit.
		(i) Behavioural signs (7)	√?				
		(ii) Paranoid symptoms (16)	√?				
		(iii) Passivity experiences (8)	√?				
		(iv) In remission or atypical delusion (13)	×				
		2. Mixed clinical (anxiety or depression) (7)	× ?				
		3. Non-clinical (40)	×				

Study	Task	Groups of participants (*N*)	Results: T.O.M. deficits by group	Results: T.O.M. deficit SCHIZ	NEG	PARA	Comments
4. Sarfati *et al.*, 1997	Non-verbal comic strip, intention inference	1. Schizophrenia (24) (i) High thought and speech disorganisation (ii) Low thought and speech disorganisation Also divided as: (a) Disorganised (5) (b) Undifferentiated (6) (c) Paranoid (7) (d) Remission (6) 2. Clinical depression (2) 3. Non-clinical (24)	√ √ × × ? × ? × ? × ? × ×	√		× ?	Performance correlated with IQ, therefore covaried out. No differences between four symptom subgroups (but very small numbers). Clear deficit in the high thought and speech disorganisation group cf other schizophrenia group, matched for total symptoms, age and IQ and cf the two control groups.
5. Langdon *et al.*, 1997	1. Picture sequencing (some T.O.M. inferences) 2. Story telling 3. Recall with T.O.M. inferences	1. Schizophrenia (20) 2. Non-clinical (20)	√ ×	√	√	× ?	The schizophrenia group showed T.O.M. deficits, but analysis of individual performance found different patterns of specific and more general cognitive deficits. Correlated analysis found no link between paranoid symptoms and specific T.O.M. deficit; instead paranoid symptoms correlated with a more general cognitive deficit (Sequencing). Negative symptoms correlated with T.O.M. and general (sequencing) deficits.
6. Doody *et al.*, 1998	1. Sally-Anne Task (first order T.O.M. inferences) 2. Ice Cream Van Task (second order T.O.M. inferences.)	1. Schizophrenia (28) 2 Psychotic depression (major affective disorder, depression (10), bipolar affective disorders, depressed (2) (12)) 3. Mild learning disability (19) 4. Mild learning disability co-morbid with schizophrenia (18) 5. Non-clinical (20)	1st order: × ; 2nd order: √ × √ √ ×	√	√	?	The relationship between T.O.M. performance and IQ and the specificity of T.O.M. deficits to people with schizophrenia were investigated. Results suggest that poor T.O.M. performance in group with schizophrenia cannot be explained by IQ alone, although IQ contributed to T.O.M. deficits. The absence of a T.O.M. deficit in the groups with psychotic depression and the non-psychiatric controls suggests that it may be specific to schizophrenia. There was no symptom subgroup analysis, however there was evidence that, in the schizophrenia group only, higher negative and higher general symptom scores were associated with poorer performance; this only approached significance for positive symptoms.

Table 1 (continued)

Study	Task	Groups of participants (*N*)	Results: T.O.M. deficits by group	Results: T.O.M. deficit SCHIZ	Results: T.O.M. deficit NEG	Results: T.O.M. deficit PARA	Comments
7. Mitchely *et al.*, 1998	1. Comprehension of irony task	1. Schizophrenia—non-acute (18) 2. Clinical control (mixed diagnoses, non-psychosis) (13)	√	√	√	?	Groups did not differ in premorbid IQ or age, but differed in current IQ and sex. Schizophrenic group showed deficit in interpretation of T.O.M. but not literal task items. Comprehension of irony (T.O.M.) correlated with current IQ in schizophrenic group, but when IQ is covaried out, group difference still found. Poor performance correlated with negative but not positive symptom scores, although paranoid symptoms not separately assessed. Authors acknowledge that validity and psychometric properties of task not established.

The other studies also used a variety of theory of mind tasks and generally employed cross-sectional group designs. Measures of IQ and of symptoms were taken. Exceptionally, Langdon, Michie, Ward, McConaghy, Catts & Coltheart (1997) used a cognitive neuropsychological approach, testing hypotheses in a series of studies, rather than the more common experimental group design. They first investigated performance on tasks and then aimed to identify associations between patterns of performance and symptom profiles. Doody, Gotz, Johnstone, Frith & Cunningham Owens (1998) investigated the contribution of IQ to theory of mind performance and the specificity of the deficit to schizophrenia by selecting groups of people with learning disability and a group with affective (depressive) psychosis.

Results from theory of mind studies

The results from all the studies showed that groups of patients with a diagnosis of schizophrenia perform theory of mind tasks more poorly than non-psychiatric controls and, in most cases, than the psychiatric controls. The deficit is greater with second order tasks, which are also generally more difficult for all participants. Two of Frith's more specific predictions are confirmed. First, the currently symptomatic patients generally perform worse than the patients in remission, the latter group showing no deficit. Secondly, patients with negative symptoms and incoherent speech are consistently shown to have a theory of mind deficit, although there is evidence of a contribution to poorer performance from a more general cognitive impairment in these patients.

However, the results are less clear in terms of Frith's predictions concerning persecutory delusions and passivity phenomena. In only two of the seven studies is there clear support for a theory of mind deficit in the subgroup with paranoid symptoms, together with the expected absence of a deficit in the passivity experiences group (Corcoran *et al*., 1995; Frith & Corcoran, 1996). This finding was not replicated by Corcoran *et al*. (1997) who found that the paranoid and passivity groups did not differ from each other and were both poorer at theory of mind tasks than normal controls, as, however, were the depressed/anxious control group. Doody *et al*. (1998) did not conduct a symptom subgroup analysis, but did find stronger evidence of an association between higher negative and general symptom scores and the deficit than with positive symptoms. Similarly Mitchley, Barber, Gray, Brooks & Livingstone (1998) found that theory of mind deficits correlated with negative, but not positive, symptoms. Two studies failed to find evidence of a specific theory of mind deficit in paranoid patients (Langdon *et al*., 1997; Sarfati, Hardy-Baylé, Besche & Widlöcher, 1997), although Sarfati *et al*.'s subgroup numbers were very small and there may also be differences between the studies in the classification of subgroups. Langdon *et al*. (1997) found correlational evidence for both a theory of mind deficit and a more general cognitive deficit in patients with negative symptoms; however, rather than a theory of mind deficit in patients with delusions, Langdon *et al*. found an unexpected but strong relationship with a more general cognitive deficit. They argue that this may be because of a deficit in deluded patients in probabilistic reasoning, with a failure to critically evaluate the plausibility of cause and effect sequences, which was consistent with their observations of patients as they

performed the tasks. It is important to note that performance on these tasks has generally been found to be clearly related to current IQ and memory, and the findings of one study are less clear when the IQ scores are taken into account (Frith & Corcoran, 1996). Doody *et al.*'s (1998) study appears to confirm that there is a contribution to poorer performance from IQ, but that there is also 'a degree of specificity of poor theory of mind performance to a diagnosis of schizophrenia which cannot be explained by the effect of IQ alone' (p. 402), a conclusion consistent with Mitchely *et al.* (1998).

Overall, from the hierarchical arrangements of the groups in Frith and colleagues' studies and the pattern of results, it is clear that the theory of mind deficit occurs to a greater extent in more symptomatic patients. Studies which demonstrate deficits in patients with schizophrenia and which are correlated with symptom severity are difficult to interpret. Do the findings merely replicate the well-established fact that patients with schizophrenia perform poorly on a variety of tasks? This is possible, although with their careful methodology and selection of subgroups, these studies do attempt to link task performance to current symptoms. The most robust finding from this set of studies is in fact that the most severely symptomatic patients and/or those with negative symptoms and/or speech and language disorder show the most striking 'other representation' deficits. Frith and colleagues acknowledge this, but argue that these patients also have a more general cognitive deficit. Langdon *et al.* (1997), in their study of how specific the deficit is, found support for this, in that there was a clear association between negative symptoms and a more general deficit.

Do people with persecutory delusions have a theory of mind deficit?

To summarise, the current evidence suggests to the present authors that deficits in understanding the mental states of others are present in patients with current symptoms of schizophrenia, and are particularly associated with negative symptoms. These co-occur with a variety of other deficits. Theory of mind deficits are also present in *some* patients with persecutory delusions. However, while theory of mind deficits may be a plausible factor in delusion formation and maintenance, the evidence linking other-representation deficits specifically to persecutory delusions is not yet convincing. Of the three approaches reviewed in this study, this is the least investigated, and its relevance to delusions needs more study.

PROBABILISTIC REASONING BIAS (GARETY)

One of the present authors, Garety, together with David Hemsley and other colleagues (e.g. Garety, 1991; Garety & Hemsley, 1994), has offered an account of delusions in general, although based on findings from studies in which the participants generally carry a diagnosis of a non-affective 'functional' psychosis, in particular schizophrenia or delusional disorder. It was argued that the literature indicates that delusions are unlikely to share a common cause but that a number of factors contribute to their formation and maintenance. Garety & Hemsley (1994) set out a multifactorial model in which past experience, affect, self-esteem and motivation play a part in some delusions, while in others biases in perception and

judgment are prominent. Normal processes of belief formation and maintenance also come into play, such as selective attention and confirmation bias. In some cases also, there is thought to be a dynamic interplay between a number of these processes which combine to operate as interacting causal mechanisms. Such an account does not generate clear predictions about group differences; rather, it demands the detailed assessment of the individual case in order to identify the activation of particular processes.

Although favouring this multifactorial account, it was proposed that the literature suggested that probabilistic reasoning might be implicated in delusions. A number of studies had identified a tendency for people with delusions to 'jump' to conclusions (see Garety & Hemsley, 1994, chapter 3). A difficulty with some of the earlier studies of reasoning and delusions, however, was the assumption that people reason logically. Studies focused on tests of formal logic, which in general were performed poorly by controls as well as people with delusions. It is clear that reasoning is not restricted to logic and that people typically employ a number of heuristic devices which guide expectations and assist reasoning, although they may also lead to errors (Ross & Anderson, 1982). In order to study how people reason, Fischhoff & Beyth-Marom (1983) have argued that a normative framework is needed to provide a conceptual framework within which the actual performance of people can be studied. This does not assume that people adhere to the standards set by the normative framework. A Bayesian model of probabilistic inference provides a useful framework for investigating probabilistic reasoning in people with delusions, since it does not simply measure valid conclusions or errors, but assesses the way conclusions are reached. Making inferences results from a combination of the strength of the prior belief and the current situational information (Alloy & Tabachnik, 1984).

Bayesian inference provides a general framework for evaluating beliefs as they are formed and maintained, since it incorporates the level of the prior belief and governs the way in which the strength of one's belief in a hypothesis should be revised in the light of new information (Fischhoff & Beyth-Marom, 1983). Garety & Hemsley predicted that, on a Bayesian inference task, people with delusions would make more rapid and overconfident judgments than other clinical and non-clinical controls. They did not hypothesize that this bias would be specific to certain types of delusions or to people with delusions only with a diagnosis of schizophrenia (Garety & Hemsley, 1994; Huq, Garety & Hemsley, 1988).

Studies of probabilistic reasoning

In a typical experiment on probability judgments (Phillips & Edwards, 1966), participants are shown pairs of containers (e.g. jars), labelled A and B, holding a large number of items such as beads of two different colours, in a particular ratio; for example, 100 beads may be divided into 85 green and 15 red, and vice versa. Participants are informed of the proportions, and the containers are removed from view. They are then told that either container is equally likely to be chosen: the initial prior probabilities in Bayesian terms are thus always 50A:50B. One of the containers is then chosen, still hidden from view, and a bead is drawn from it and shown to the participant. The experiment is continued, with beads being drawn sequentially and

Table 2. Bayesian probabilistic reasoning studies

Study	Groups of participants (*N*)	Task materials/ratio	Results: Jumping to conclusions: Draws to decision	Probability estimates	Disconf bias	Errors	Comments
1. Huq *et al*., 1998	1. Deluded, schizophrenia (15) 2. Mixed psychiatric control (10) 3. Non-clinical control (15)	Coloured beads/85:15	combined means across conditions √		N/A	NS	The data from conditions 1 or 2 were combined so that JTC is a composite of draws to decision and of probability estimates. JTC found in deluded group relative to both control groups. Groups did not differ in IQ.
2. Garety *et al*., 1991	1. Deluded schizophrenia (13) 2. Deluded, delusional disorder (14) 3. Anxious control (14) 4. Non-clinical control (13)	Coloured beads/85:15	√	×	√	NS	No differences between delusional groups, who both showed JTC on condition 1, cf a non-clinical control group; only schizophrenic group also differed from anxious group. Disconfirmatory bias in both deluded groups. Errors: non-significant, but higher in deluded groups. Groups did not differ in IQ, but JTC associated with lower IQ and current hallucinations.
3. Mortimer *et al*., 1996	1. Schizophrenia (43)	Coloured beads/85:15	√	N/A	N/A	N/A	Correlational study. Not all participants deluded. Number currently deluded not given. Significant cognitive impairment (rehabilitation sample). 42% of sample reached decision after one draw. JTC not correlated with 'index of deludedness' (severity and number of delusions) or with IQ.
4. Dudley *et al*., 1997*a*	1. Deluded schizophrenia (15) 2. Depressed control (15) 3. Non-clinical control (15)	Coloured beads/85:15; 60:40	√	N/A	×	?	JTC in both 85:15 and 60:40 ratios in deluded cf both control groups. Results did not differ in deluded group with memory aid. Slightly increased error rates in deluded group—significance not tested. Groups did not differ in IQ or age.
5. Dudley *et al*., 1997*b*	1. Deluded schizophrenia (persecutory and grandiose) (15) 2. Depressed control (16) 3. Non-clinical control (15)	(i) concrete emotionally neutral; (ii) emotionally salient/60:40	√	N/A	N/A	?	JTC found in deluded group with both concrete emotionally neutral and emotionally salient material cf both control groups. More errors in deluded group and more with salient material (significance not tested). Emotional salience reduced data gathering in all groups; possibly greater reduction in deluded group. Groups did not differ in IQ (low IQ excluded).
6. Peters *et al*., 1997	1. Deluded mixed diagnoses (23) 2. Depressed control (22) 3. Non-clinical control (36)	Coloured beads/85:15	√	√	√?	?	JTC found in deluded group cf both control groups, in both condition 1 (draws to decision) and condition 2 (higher early probability estimates). Disconfirmatory bias in deluded group relative to non-clinical controls but not to depressed controls.

Study	Groups of participants (*N*)	Task materials/ratio	Results: Jumping to conclusions: Draws to decision	Probability estimates	Disconf bias	Errors	Comments
7. Young & Bentall, 1997*a*	1. Deluded, persecutory (schizophrenia and delusional disorder) (19) 2. Depressed controls (19) 3. Non-clinical controls (19)	Study (i): Coloured beads/60:40; 75:25; 90:10; 80:20 Study (ii): Coloured beads/85:15; and emotionally salient/85:15	N/A	×	√?	√	No test of draws to decision. Results calculated as mean scores of blocks of five estimates. In general, no differences between groups in probability estimates, for Study 1. In Study 2 both deluded and depressed clinical groups higher mean initial certainty than non-clinical controls and fewer draws to certainty (no difference between deluded and depressed). Disconfirmatory bias in deluded group relative to non-clinical controls but not to depressed controls Differences found in all three groups with emotionally salient material—faster to initial certainty and faster to change estimates. Groups did not differ in IQ or age.
8. Fear & Healy, 1997	1. Deluded, delusional disorder (30) 2. Obsessive compulsive disorder (29) 3. Mixed obsessional/deluded (16) 4. Non-clinical controls (30)	Coloured beads/85:15	√	×	√?	√?	In condition 1, JTC high proportion of deluded delusional disorder group (73%) differed significantly from other three groups. In condition 1 delusional disorder group significantly more errors than non-clinical controls, but not than clinical control groups. In condition 2, delusional disorder group greater disconfirmatory bias than controls, but not than OCD group. Group 3 not all actively deluded. Groups did not differ in IQ, age or sex.

always replaced. Although the participants are told that beads are being selected randomly, the sequence of colours is predetermined according to the ratio of the two colours. The task is to work out whether the experimenter is drawing from container A or container B. Typically, the experiment has two conditions: 'draws to decision' and 'probability estimates'. In 'draws to decision', the participants are free to determine how many beads are drawn and the trial is only terminated once they affirm that they are certain about their choice. In 'probability estimates', participants are asked to indicate at each stage in the sequence estimates of the probability of one container having been chosen rather than the other. In this condition, there is a fixed number of trials. There is an 'optimal' strategy for performance on these tasks which can be computed using Bayes' theorem. It has been noted that the general population is consistently conservative in these tasks, requiring more draws than Bayes' theorem would indicate to reach a decision and giving under-confident probability estimates (Edwards, 1982).

Eight studies of probabilistic reasoning, all using modifications of the basic paradigm, are shown in Table 2. In most studies coloured beads are used in ratios of 85:15. This ratio creates an easy task, which minimises floor effects. In addition to investigating jumping to conclusions (JTC) on the basis of either draws to decision or probability estimates, Garety, Hemsley & Wessely (1991) and four subsequent studies examined the readiness of participants to switch their hypotheses when given potentially contradictory evidence. Dudley, John, Young & Over (1997*a*,*b*) and Young & Bentall (1997*a*) investigated performance with different ratios of beads; they also introduced concrete and emotionally salient task materials. There are some other differences between studies: these are in the selection criteria for people with delusions (diagnosis; types of delusions), the selection of clinical control groups, and the total number of trials given.

Results from probabilistic reasoning studies

The main findings from these studies can be summarised thus. First, JTC has been found in seven studies. Only Young & Bentall (1997*a*) failed to replicate the finding. JTC was found in all seven studies using draws to decision. In contrast, JTC is *not* found when a fixed number of trials are presented and probability estimates are required; in those cases, the deluded participants performed more like the controls, although two studies (Huq *et al*., 1988; Peters, Day & Garety, 1997) found some differences relative to controls. JTC is not a function of a memory deficit, since results were unchanged by the presence of a memory aid, or of impulsiveness, since the deluded group adjusted the amount of evidence required with a changed probability ratio (Dudley *et al*., 1997*a*). JTC is found in people with delusions, irrespective of a diagnosis of schizophrenia or of delusional disorder, and in people with schizophrenia where the current delusional status is less clear (Mortimer *et al*., 1996). Mortimer *et al*. did not find that JTC correlated with the number or severity of current delusions (on an 'index of deludedness'). The studies show a higher error rate in deluded participants, although it is in absolute terms low and only statistically significant in two studies (Fear & Healy, 1997; Young & Bentall, 1997*a*); there is a further suggestion that errors are more numerous with emotionally salient material.

There are a number of variations in the findings. Most studies have employed both a clinical and a non-clinical control group. The finding of JTC always differentiated the deluded groups from the non-clinical, the depressed and the OCD controls; however, in the Garety *et al*. (1991) study an anxious control group did not significantly differ from the delusional disorder group, although they did differ from the schizophrenia group. JTC is more likely across groups for emotionally salient material, but this may be further exaggerated in people with delusions (Dudley *et al*., 1997*b*). A disconfirmatory bias (i.e. shifting probability estimates downwards following potentially disconfirmatory information) was found in four studies when deluded groups were compared to non-clinical controls (Fear & Healy, 1997; Garety *et al*., 1991; Peters *et al*., 1997; Young & Bentall, 1997*a*); however, this did not always distinguish deluded groups from depressed and OCD controls.

Although groups of deluded participants typically had a higher proportion of males, there is no evidence of a sex difference. In general, the groups did not differ in IQ; however, the contribution of IQ to performance has not been systematically studied. Garety *et al*. (1991), on a *post hoc* analysis, found evidence that JTC was associated with lower IQ, while Mortimer *et al*. (1996) did not find a correlation with IQ, as measured by the mini mental state examination. Finally, the most extreme reasoning bias (using only one item of information, on an 85:15 ratio) is not present in all people who are currently deluded, although it is present in a significant proportion (where given, percentages range between 40 and 70%).

Maher (1992) has debated the interpretation of the findings, referring specifically to the first (Huq *et al*., 1988) study. He points out that the difference between the groups does not indicate that the people with delusions showed faulty inferences: the participants with delusions made few errors and the mean number of draws taken to reach a decision (2.2) represents *better* Bayesian reasoning than the control groups. After two beads of the same colour have been drawn, the objective Bayesian probability that the beads have been drawn from the jar with predominantly beads of that colour is 97%. This, Maher (1992) argues, was therefore a reasonable point to make a decision. The control subjects, as has been previously found, were in contrast over-cautious in their estimates (Edwards, 1982). Thus, although the results indicate that the people with delusions request less information before reaching a decision than controls, Maher (1992) argues that the findings fail to demonstrate any support for 'the faulty-inference hypothesis of schizophrenic delusions' (p. 106).

It is noteworthy that, as a group, the people with delusions responded according to Bayesian norms, in contrast to the conservative performance of controls. However, there was considerable variability in the deluded group, and nearly half made a decision on the basis of one draw (rather than two), at an objective probability of 85%, which is very rare in controls and can be considered incautious, when certainty is the task criterion. These patterns were repeated in subsequent studies. Overall, these findings should not be interpreted as evidence of a deficit, an *inability* to reason probabilistically or to test hypotheses, but rather of a tendency or bias to the early acceptance and, to a lesser extent, the early rejection of hypotheses. This, may, under certain conditions, contribute to erroneous inferences and, therefore, to delusion formation.

Table 3. Further reasoning studies

Study	Groups of participants (*N*)	Task	Results: Reasoning bias in deluded participants	Results: Comments
1. John & Dodgson, 1994	1. Deluded, schizophrenia (12) 2. Depressed (12) 3. Non-clinical (12)	20 Questions game (inductive reasoning, information gathering to reach conclusion)	√	Deluded subjects requested less information than either control group before reaching a decision and produced poorer judgments. Findings are independent of age, years in education or sex.
2. Young & Bentall, 1996	1. Persecutory delusions (schizophrenia or delusional disorder (16) 2. Depression (16) 3. Non-clinical (16)	Hypothesis testing (visual discrimination problems, discovering rules for correct choices, with feedback given at pre-determined points)	√	Participant groups predominantly male (only one female in each group) and matched for years in education and performance IQ. Trend for deluded group to make fewer correct hypotheses than either control group. Deluded participants less inclined to stick to hypotheses when given positive feedback, with evidence to suggest relationship between task performance and non-verbal IQ. Results thought to reflect basic deficits in ability to make use of sequential data.
4. Bentall & Young, 1996	1. Persecutory delusions (Schizophrenia or delusional disorder (15) 2. Depression (15) 3. Non-clinical (15)	Sensible hypothesis testing (questionnaire measure, choice between three methods of testing hypothesis concerning a positive or a negative outcome).	×	Groups matched for age, IQ, sex and years full-time education. No differences between groups, and 'sensible reasoning' strategies were replicated in all groups. Findings suggest people with delusions do not have difficulty in deciding how to go about testing hypotheses.
4. Kemp *et al*., 1997	1. Deluded (schizophrenia, delusional disorder and atypical) (16) 2. Non-clinical (16)	Formal logic reasoning tasks (conditionals, syllogisms and probability), emotional content varied.	√	There are methodological problems in that there was no clinical control group, IQ was not assessed in non-clinical control group and groups were not matched for age or sex. Some evidence of a relationship between cognitive deficit (cognitive estimates)/age and task performance in deluded group. Overall, both groups showed poor performance on these tasks; however deluded group did differ in endorsing more invalid or fallacious responses, especially when emotive themes are involved.

Study	Groups of participants (*N*)	Task	Results: Reasoning bias in deluded participants	Results: Comments
5. Dudley *et al*., 1997*a*	1. Delusions (persecutory/grandiose) mixed diagnoses (12) 2. Depression (12) 3. Non-clinical (12)	Probability estimates (biased coin task)	×	No differences between the groups in age or IQ (current and estimated pre-morbid). There were no differences between the groups on task performance, and all groups altered probability estimates in the same way in response to changes in the strength of the evidence. The authors conclude that the apparent discrepancy with previous probabilistic reasoning research may be owing to an important task difference—in this study, participants are given all the information to be considered rather than being free to decide when they have sufficient information.
6. Linney *et al*., 1998	General (student) population (40) divided into: 1. High delusional ideation (20) 2. Low delusional ideation (20)	Four tasks—hypothesis testing: 1. Wason's 2–4–6 problem 2. Wason's selection task and probability judgment: 3. Coin tossing task 4. Book/suicide problem	√*	No differences between groups in age, years of education, IQ and social desirability questionnaire scores. Differences found between groups on two tasks (one hypothesis testing—modified 2–4–6 problem; one probability judgment—coin-tossing task) but not on other two tasks involving hypothesis testing and aggregation of probabilistic information. Differences reflected 'jumping to conclusions' on 2–4–6 task and taking less account of varying sample sizes on coin tossing task. Authors conclude that there is evidence of a data-gathering bias, independent of whether information is pre-determined. Because this is found in a non-clinical high delusional ideation population, they argue it may suggest that a reasoning bias is implicated in delusion formation rather than just delusion maintenance.

* high delusional ideation

Evidence of reasoning bias from studies employing other tasks

The above studies showed that, on a particular test of probabilistic reasoning, people with delusions jumped to conclusions. However, it is not clear how specific the findings are to this reasoning task. A robust test of the hypothesis of rapid and over-confident judgments in people with delusions involves examining performance on other tasks. There have been six further studies recently conducted to explore reasoning in people with delusions further. All of these studies explicitly aimed to build on Garety and colleagues' work by employing different tasks and normative frameworks to consider how general a bias there may be and/or to specify more precisely the biased reasoning processes in people with delusions. The studies have examined probability estimation, hypothesis testing, inductive reasoning and data gathering and formal logic (see Table 3). All the studies (except Linney, Peters & Ayton, 1998) recruited currently deluded participants, while Linney *et al*. recruited participants from the general population and divided them by high and low delusional ideation. Clinical control groups were generally drawn from populations of depressed patients, although Kemp, Chua, McKenna & David (1997) did not include such a control group.

The findings of these studies are mixed. Four studies found statistically significant differences in performance between the deluded and control groups (John & Dodgson, 1994; Kemp *et al*., 1997; Linney *et al*., 1998; Young & Bentall, 1995), while two of these studies found no differences (Bentall & Young, 1996; Dudley *et al*., 1997*a*). Taken together, these negative and positive findings provide interesting pointers to the nature of the reasoning bias in people with delusions. Bentall & Young's (1996) study suggests that people with delusions are able to test hypotheses when supplied with a range of options, while Dudley *et al*. (1997) found that people with delusions are similar to controls in their estimation of probabilities, again when all the information is supplied. The differences found by Young & Bentall (1995) and Kemp *et al*. (1997) represented poorer performance in both studies and were thought by the authors to reflect a difficulty making use of sequential information and a weakness in formal logic respectively. The findings of both John & Dodgson (1994) and Linney *et al*. (1998) were interpreted as reflecting differences in gathering information. In John & Dodgson, the deluded participants demonstrated a 'difference in cognitive processing style which limits the extent to which deluded subjects request information to help them form a decision' (p. 45). Linney *et al*. studied the reasoning performance on four tasks of participants in the general population who were high in delusional ideation. No differences were found on tasks involving hypothesis testing and the aggregation of probabilistic information, indicating that the bias may not reflect difficulties with sequential information, contrary to Young & Bentall's (1995) speculation. There were, however, differences on two tasks which the authors interpreted as reflecting a JTC style of data gathering. This bias in data gathering was independent of whether the information provided was pre-determined.

In these studies, the groups were generally matched for IQ and age and there is little to suggest that performance varied with IQ, excepting the study of formal logic by Kemp *et al*. (1997). However, the use of multivariate analyses, in which IQ is

covaried, would provide a better test of the involvement of IQ. Linney *et al*.'s (1998) study is the only one to employ a non-clinical population divided by delusional ideation. That there are differences in a non-clinical sample may suggest that the data-gathering bias is implicated in delusion formation, although clearly it would be informative to attempt to replicate these performance differences both in another non-clinical population as well as in a clinical population.

Do people with delusions have reasoning bias?

Of the 14 studies reviewed, 11 provide evidence for reasoning biases in people with delusions. The clearest findings derive from the studies employing the Bayesian probabilistic reasoning paradigm: all seven studies which investigated data gathering found evidence of JTC. Taking all the studies together, a picture emerges of people with delusions showing a tendency to seek less information to reach a decision, but not, when presented with information, being unable to use it. People with delusions do not, it seems, have a probabilistic reasoning bias, as Garety and colleagues have previously suggested, in that they can estimate probabilities, but have a data-gathering bias. This willingness to accept a hypothesis on the basis of less evidence than control groups is supported by two studies employing different paradigms (John & Dodgson, 1994; Linney *et al*., 1998). Contrary to Maher's hypothesis, it does seem that the reasoning bias found in people with delusions can lead to the acceptance of incorrect hypotheses. Moreover, the 'disconfirmation' bias found in a number of the Bayesian studies suggests that people with delusions may be more ready to abandon existing hypotheses and form new ones, again on the basis of little evidence. These findings apply to 'neutral' task material, where the content is not thought to be of particular relevance to participants' concerns or to be emotional in content. However, there is evidence that emotional salience affects people's reasoning in general and possibly affects the reasoning of people with delusions to a greater extent (Dudley *et al*., 1997*b*; Kemp *et al*., 1997; Young & Bentall, 1997*a*). The negative findings of these studies are also informative. They show that the results do not simply reflect deficits in performance but can be specified with some precision. The evidence that people with delusions are competent at certain reasoning tasks, such as testing hypotheses and aggregating information, is important clinically as well as theoretically.

PERSECUTORY DELUSIONS AS DEFENCE (BENTALL)

Bentall and his colleagues have proposed that people with persecutory delusions construct them to maintain self-esteem, avoiding discrepancies entering consciousness between how they perceive themselves to be and how they would like to be (Bentall, Kaney & Dewey, 1991; Bentall, Kinderman & Kaney, 1994; Kinderman & Bentall, 1996*a*). They argue that externalising causal attributions (persecutory delusions) are evoked for negative events which might otherwise increase the accessibility of underlying negative self-representations. As these researchers acknowledge, these ideas are similar to those of earlier theorists. For example, Zigler & Glick (1988) suggested that paranoia is a form of camouflaged depression and Colby, Faught & Parkinson (1979) explained persecutory delusions as the product of

a tendency to perceive threat to self-esteem combined with a protective externalising attributional bias. Bentall and colleagues' account is largely based on two sets of experimental findings. The first set concerns attributional style; the second concerns discrepancies between overt and covert self-esteem.

Studies of attributional biases

Attributional style has been investigated primarily using the Attributional Style Questionnaire (ASQ; Peterson, Semmel, Von Baeyer, Abramson, Metalsky & Seligman, 1982), which was originally used to investigate cognitive processes in depression. In the ASQ, a positive or a negative event involving the respondent is briefly described (e.g. 'You get a pay rise'). The respondent is instructed to write down one major cause for the event, and then to rate the self-generated cause on three dimensions: internality, stability and globalness. It is the internality dimension (the cause is 'due to something about you or something about other people or circumstances') which is of central importance to the persecutory delusion-as-defence account, since it is proposed that an attributional style characterised by blaming other people for bad events and taking credit for good events is implicated in persecutory delusions. Bentall and colleagues describe this as an extreme form of the self-serving bias, which is reported to be present in the general population as a means of maintaining self-esteem (Bentall *et al*., 1994). There are two methods of treating the internality data: one method involves calculating a composite difference score, which consists of attributional style for positive events minus attributional style for negative events and treating this measure as a single internality construct; the other method treats attributions for positive and negative events separately. Byrne & MacLeod (1997) criticise the former practice on the grounds that attributional styles for positive and negative events show a low degree of correlation and argue that attributions for positive and negative events should be treated separately. The present authors have therefore favoured this second method in reporting results in Table 4, noting where the reported composite results do not permit inferences separately on internality for positive and negative events.

The ASQ has been criticised for poor reliability, in particular of the internality dimension, and Kinderman & Bentall (1996*b*) have consequently developed a new scale, the Internal Personal and Situational Attributions Questionnaire (IPSAQ). They propose that three distinct attributional loci can be identified on the ASQ internality dimension: an internal locus (attributing events to oneself), an external–personal locus (attributing the event to identifiable others), and an external–situational locus (attributing the causes of events to situations or chance). For example, if the negative event for which an explanation is requested is 'You are late for an appointment', it is possible to attribute this externally either to a person (my partner made me late) or to a non-personal situation (the rain made the traffic terrible). Finally, there is a third questionnaire measure of attributions, the Social Attributions Questionnaire (SAQ) constructed by Bentall *et al*. (1991), which is concerned with judgments about the social interactions between two other persons. Attributions are made to the actor (a person attribution), the target or victim (a stimulus attribution) or to the circumstances.

Table 4. Studies of attributional biases

Study	Groups of participants (*N*)	Measure	Results: Externalising bias for negative events	Results: Internalising bias for positive events	Comments
1. Kaney & Bentall, 1989	1. Persecutory delusions (paranoid schizophrenia and delusional disorder) (17) 2. Depressed or other psychiatric disorder (16) 3. Non-clinical (197)	ASQ	√?	√?	Internalising bias *relative* to externalising bias, i.e. composite positive–negative events score reported as significant. Visual examination of the graphs suggests this accounted for more by high externality for negative events than high internality for positive events.
2. Candido & Romney, 1990	1. Paranoid disorder or paranoid schizophrenia (15) 2. Paranoid disorder or paranoid schizophrenia and depression (15) 3. Depression (15)	ASQ	√?	√?	No non-clinical control group—biases found are relative to a depressed sample who show the typical depressive bias (internal for negative events and external for positive events). Findings suggest it is important to examine role of concurrent depression in attributional style.
3. Lyon *et al*., 1994	1. Persecutory delusions (paranoid schizophrenia and delusional disorder) (14) 2. Depressed (14) 3. Non-clinical (14)	ASQ (pf)	√	×	Although 'self-serving bias' composite score is reported, it is accounted for by high externality for negative events, not high internality for positive events, relative to controls—on basis of visual examination of graphs (also personal comm (Bentall)).
4. Fear *et al*., 1996	1. Delusional disorder (29): (i) persecutory (20) (ii) non-persecutory (9) 2. Non-clinical (20)	ASQ	√	×	Delusional disorder group more external and stable for negative events. No difference between persecutory delusions and non-persecutory delusions (although the latter includes five with grandiose delusions and group *N* is small). Delusional disorder not more internal for positive events.
5. Sharp *et al*., 1997	1. Delusional disorder (31) (i) Persecutory/grandiose delusions (19) (ii) Non-persecutory/non-grandiose delusions (12) 2. Non-clinical (24)	ASQ	√	×	Persecutory delusions subgroup more external for negative events than non-persecutory delusions group and than controls. Non-persecutory delusions group did not differ from controls. Attributional style may shape delusional content rather than delusional form.

Table 4 (continued)

Study	Groups of participants(*N*)	Measure	Externalising bias for negative events	Person attributions for negative events	Person attributions for positive events	Comments
6. Kinderman & Bentall, 1997	1. Persecutory delusions (schizophrenia 18) or delusional disorder (2)) (20) 2. Depressed (20) 3. Non-clinical (20)	IPSAQ	×	√	×	Groups approximately matched for age and sex and the patient groups did not differ in duration of illness. The depressed group made significantly more internal attributions for negative events than both other groups, while both paranoid and non-clinical controls made significantly more internal attributions for positive than negative events than the depressed group. The paranoid group made significantly more person attributions for negative events than both control groups but did not differ in person attributions for positive events.
7. Bentall *et al*., 1991	1. Persecutory delusions (schizophrenia (4) or delusional disorder (13)) (17) 2. Depressed (depression (13), anorexia (2) or schizophrenia (2)) (17)	SAQ		√		Groups approximately matched for age, years of education and sex. Psychometric properties of SAQ not established. Deluded participants made significantly more person attributions than both control groups, and more person attributions for negative actions. Deluded group were significantly more certain of their judgments than depressed, but not than non-clinical controls.
8. Young & Bentall, 1997*b*	1. Persecutory delusions (schizophrenia (10) or delusional disorder (6)) (16) 2. Depressed (16) 3. Non-clinical (16)	SAQ		×		No differences between the groups which all made attributions and changed attributions in directions predicted by attribution theory. Some differences in task method from above study. Groups did not differ in sex and estimated IQ. Authors suggest deluded participants may only make inflexible judgments with self-referent materials.

ASQ: Attributional Style Questionnaire (pf: parallel form).
IPSAQ: Internal, Personal and Situational Attributions Questionnaire.
SAQ: Social Attributions Questionnaire.

On the ASQ, Bentall and colleagues predict that people with persecutory delusions will make externalising attributions for bad events and internalising attributions for good events, relative to controls. On the IPSAQ, they predict that paranoid participants will make excessively external attributions for bad events relative to depressed but not non-clinical controls, and that paranoid participants will make more personal rather than situational attributions relative to both groups of controls. Finally, for SAQ, they predict that patients with persecutory delusions will make more person attributions than controls. They do not make any predictions about non-persecutory delusions and they do not make predictions concerning diagnostic groups.

Most of the studies shown in Table 4 have been conducted by Bentall, Kaney, Kinderman and co-workers, although two other groups have contributed three studies (Candido & Romney, 1990; Fear, Sharp & Healy, 1996; Sharp, Fear & Healy, 1997). All the studies employ a cross-sectional group design. Bentall and colleagues have adopted the general strategy of comparing a group of people with persecutory delusions with two control groups: people who are depressed and a non-clinical sample. In most cases, the persecutory delusions group consists of people with a diagnosis of schizophrenia (the majority) or of delusional disorder; their other current psychotic symptoms are not generally reported, although depression scores are. Studies typically have more male than female participants.

Results of attributional bias studies

Table 4 shows five studies which have employed the ASQ, and whether an externalising bias for negative events and an internalising bias for positive events were found. It shows clear evidence that people with persecutory delusions, when compared to depressed and non-clinical control groups, show a bias to excessively external attributions for negative events. All five studies support this conclusion. The evidence in support of an internalising bias for good events is much less strong. Three out of five studies did not find this bias (Fear *et al*., 1996; Lyon, Kaney & Bentall, 1994; Sharp *et al*., 1997). The study by Candido & Romney (1990) only provides a comparison with a depressed group (well known to demonstrate the opposite bias) and no non-clinical control, and the study by Kaney & Bentall (1989) reported a composite analysis of attributional style, as described above, so that the results are difficult to interpret. Taking the results of these five studies together, it cannot therefore be concluded that people with delusions are excessively self-serving: in general they do not differ from non-clinical controls in taking credit for good events.

Using the IPSAQ, Kinderman & Bentall (1997) found that the participants with persecutory delusions showed a personalising bias for negative events (that is they blamed other people for them) in comparison to both depressed and non-clinical controls. They were, however, not more likely to show either an externalising bias for negative events or an internalising bias for positive events than the non-clinical controls: that is, overall, the patients with delusions did not attribute blame externally for bad events or take credit for good events. It therefore seems probable that the attributional bias of people with persecutory delusions is not in general

externalising or specifically self-serving, but rather personalising: a tendency to blame other people when things go wrong.

The results of two further studies of attributions in people with delusions, using the SAQ, are also shown in Table 4 (Bentall, Kaney & Dewey, 1991; Young & Bentall, 1997*b*). The first study found a tendency for people with delusions to make person attributions, compared to controls, for negative events; however, recently, using a somewhat different methodology, Young & Bentall (1997*b*) failed to replicate this. Indeed, Young & Bentall found no difference at all between groups in their attributions and in changes made to attributions when new information was given. They concluded that self-referent tasks may be essential for demonstrating attributional biases in patients with persecutory delusions.

Fear *et al*. (1996) and Sharp *et al*. (1997) have investigated the specificity of an externalising attributional style to persecutory delusions, focusing their investigations on people with a diagnosis of delusional disorder. They thus address two questions of specificity: are attributional biases found in people with delusional disorder, since most previous studies were primarily of people with a diagnosis of schizophrenia, and are these biases confined to persecutory delusions? The findings are mixed. There are preliminary grounds for hypothesising that the externalising bias for bad events may be specific to persecutory and grandiose delusions. Sharp *et al*. (1997) conclude that their findings argue against a primary aetiological role for attributional style in the genesis of delusions, but suggest that such a style may shape delusional content.

Are attributions biased in people with persecutory delusions?

In terms of the key question of attributional style, a fairly consistent picture emerges from these eight studies. People with persecutory delusions do have attributional biases: they show an externalising bias for bad events, when the material is self-referent. When these external attributions are further divided into personalising and situational, there is early evidence that people with delusions are particularly inclined to blame people rather than situations or chance; when attributions are examined in this way, they may not be more generally inclined to externalise blame. On the current evidence, people with persecutory delusions do not seem to be particularly biased to internalising when attributing causes for good events: overall, then, they are not particularly self-serving. In sum, people with persecutory delusions, when confronted with self-referent material, are particularly likely to see other people as responsible for bad events.

Self-discrepancies

The second set of experimental evidence, invoked to support the proposal that persecutory delusions act as a defence, concerns studies of discrepancies between overt and covert levels of self-esteem or self-representations. This proposal is complex, both conceptually and methodologically, and has been developed over time. Therefore the following first briefly considers both the conceptual issue, which concerns the proposal of discrepancies in self-esteem and the self-concept, and the

methodological problem of eliciting such discrepancies, before reviewing the relevant studies.

Self-esteem is an evaluative component of the 'self-concept', a broader term which embraces beliefs concerning relatively specific aspects of the self (such as musical ability, gender identity and so on) and self-esteem (Brewin, 1988). In earlier formulations of their theory, Bentall and colleagues argued that persecutory delusions reflect a defensive, self-serving attributional style which protects the individual against underlying feelings of low self-esteem (e.g. Bentall, 1994; Lyon *et al*., 1994). Bentall *et al*. (1994) have adopted a more elaborate model proposed by Higgins (1987) which invokes 'self-discrepancies' in terms of three basic domains (actual self, ideal self and ought self) and in terms of two basic standpoints: the point of view of the self and the viewpoint of the other. Higgins argues that the term 'low self-esteem' generally refers either to a negative actual self-concept or to a discrepancy between the actual self-ideal self. Different forms of emotional discomfort (depression, anxiety, shame) are predicted by specific self-discrepancies. For example, dejection-related emotions (dissatisfaction, sadness) are predicted by the presence of an actual self-ideal self-discrepancy.

Bentall *et al*. (1994), drawing on Higgins' account, have refined their proposal, positing that the person with persecutory delusions has an attributional bias which serves to minimise discrepancies between the actual self and the ideal self at the cost of externalising causal attributions and thereby generating discrepancies between the actual-self and the actual-other domains. Bentall and colleagues do not generally claim that the defensive function of persecutory delusions protects the individual from depression, although it seems that this should be predicted by Higgins' model since it is the actual self-ideal self-discrepancy which is associated with dejection-related emotions. They propose that the self-serving attributional bias minimises the awareness of discrepancies between actual–ideal representations of the self, in order to maintain a positive explicit self-concept.

Thus Kinderman & Bentall (1996*a*) predicted that paranoid delusions would be associated with positive self-ratings in the actual-self domain and a high degree of consistency between the actual self-concept and both the ideal-self and ought-self concepts. Similarly, Bentall (1994) noted that a paradoxical combination of 'high self-esteem and high depression' will occur if delusions function to defend against low self-esteem. The model predicts that people with persecutory delusions will show overt normal or high self-esteem, or a positive explicit self-concept. However, in a 'weaker' formulation of the delusion-as-defence account, it could be argued that the delusion is only partially successful (i.e. it does not fully preserve self-esteem). Therefore the most important prediction of the model is that there will be a discrepancy between overt and covert measures of self-esteem, in that overt measures will not show actual–ideal self-discrepancies, while on covert measures such a discrepancy will be detected.

Studies of overt and covert negative self-concepts: detecting a discrepancy

The studies reviewed have used different methods to assess overt self-esteem (see Table 5). First, there are standardised questionnaire measures, the Coopersmith Self-

Table 5. Studies of overt self-esteem, covert self-esteem and of overt-covert discrepancies

Study	Groups of participants (*N*)	Measure/ task	Results: Overt normal/high self-esteem	Results: Covert low self-esteem	Results: Discrepancy	Comments
1. Candido & Romney, 1990 (see also Table 4)	1. Paranoid disorder or paranoid schizophrenia (15) 2. Paranoid disorder or paranoid schizophrenia and depression (15) 3. Depression (15)	CSEI	√ ×	N/A	N/A	Groups differ from each other significantly. Note the absence of a non-clinical control group. Although not given in the paper, the norms for the CSEI show that the non-depressed paranoid group have normal self-esteem, while the depressed paranoid group have low self-esteem (more than 1 SD below mean) and the depressed group very low self-esteem (more than 2 SD below mean). Thus, self-esteem scores decreased as depression increased, although no correlation is reported. The study does not make clear whether patients in the 'paranoid' groups all had persecutory delusions.
2. Lyon *et al*., 1994 (see also Table 4)	1. Persecutory delusions (paranoid schizophrenia and delusional disorder (14) 2. Depressed (14) 3. Non-clinical (14)	1. RSEQ 2. ASQ pf 3. PIT	√	√	√	Groups approximately matched for age, sex and intelligence. The deluded group did not differ from the non-clinical control group in self-esteem and both groups had significantly higher self-esteem than the depressed group. The authors state that the deluded group exhibited a profile of 'high self-esteem and high depression' (p. 640). Correlations between depression and self-esteem were negative and significant in the deluded ($r = -.49$) and non-clinical control groups ($r = -.54$). Study also tested hypothesis of a defensive attributional style protecting against harboured feelings of low self-esteem. A discrepancy was found between overt measure (ASQpf) which found excessively external attributions for negative events relative to both control groups and covert PIT, which found that the persecutory delusion group, like the depressed group, made internal attributions for negative events and external attributions for positive events.
3. Kinderman & Bentall, 1996*a*	1. Persecutory delusions (schizophrenia (18) and delusional disorder (2)) (22) 2. Depression (22) 3. Non-clinical (22)	PQQ	√	N/A	N/A	Groups were approximately matched for age and gender and the clinical groups for duration of illness. Using a measure of self-representations, actual and ideal and ought self-representations found to be consistent in paranoid patients but marked discrepancies were found between self-perceptions and believed perceptions of parents about the self. Paranoid patients did not differ from non-patient controls in self-representations, but differed from depressed patients who showed expected self-discrepancies in actual—ideal self-domains. Persecutory group was moderately depressed, but relationship between self-representations and depression was not analysed.

Study	Groups of participants (*N*)	Measure/ task	Results: Overt normal/high self-esteem	Results: Covert low self-esteem	Results: Discrepancy	Comments
4. Freeman *et al*., 1998	1. Persecutory delusions (28) 2. Other positive symptoms of psychosis (25)	SCQ	×	N/A	N/A	This study used data from a therapy trial. Data were analysed cross-sectionally at baseline, and longitudinally separately for the therapy and treatment as usual control. It examined levels of self-esteem in patients with and without persecutory delusions and the relationship of self-esteem to measures of depression and of delusions. Both groups showed low self-esteem relative to population norms. Negative correlations were found between depression and self-esteem, both cross-selectionally and longitudinally in people with persecutory delusions. About 30% of people with persecutory delusions had normal self-esteem, and a cross-sectional analysis found higher conviction in their delusions and lower depression and anxiety. Although not replicated longitudinally, the authors conclude this may be consistent with a defensive process for a subgroup.
5. Kinderman, 1994	1. Persecutory delusions (16) 2. Depressed (16) 3. Non-clinical controls (16)	1. PPQ 2. EST	× ?	√	?	This study was a direct test of discrepancies between overt self-ratings (as self-descriptive, 30 positive and 30 negative personally descriptive adjectives) and the covertly accessed responses of the EST (using a sample of the same adjectives). The persecutory group was mildly/moderately depressed. On the EST, both patient groups were slowed by negative words, while persecutory group only was slowed relative to non-clinical controls by positive words. This is consistent with delusional disorder group having a negative self-concept, which might be related to depression. The overt ratings are difficult to interpretation (see text) but the persecutory group did not differ from depressed group in extent to which they endorsed as self-descriptive negative adjectives. Therefore a discrepancy does not seem to be found.
6. Bentall & Kaney, 1989	1. Persecutory delusions (schizophrenia or delusional disorder) (16) 2. Depressed (eating disorders or depression) (16) 3. Non-clinical control (16)	EST	N/A	×	N/A	The study was designed to investigate attentional biases rather than self-discrepancies. Both persecutory and depressed groups were depressed. While persecutory group was specifically slowed in colour naming threat content words, the depressed group was selectively slowed in colour naming depressive content words. This study does not support the hypothesis that the persecutory group had an implicit negative self-concept.

Table 5. (continued)

Study	Groups of participants (*N*)	Measure/ task	Results: Overt normal/high self-esteem	Results: Covert low self-esteem	Results: Discrepancy	Comments
7. Fear *et al*., 1996 (see also Table 4)	1. Delusional disorder (29) (i) Persecutory (20) (ii) Non-persecutory (9) 2. Non-clinical control (20)	1. EST 2. ASQ re-rating 3. DAS	N/A	×	N/A	The EST part of the study largely designed as a replication of Kaney & Bentall (1989) with delusional disorder patients and with addition of anxiety words. The EST was employed to identify 'covert depressive biases'. Both persecutory and non-persecutory group slowed by colour naming threat words, while non-persecutory group *only* differed from controls in slowing with depression and anxiety-related words. Authors conclude that there is an inferred absence of covert depressive cognitions in persecutory group. The study failed to replicate Kinderman *et al*. (1992) on the ASQ re-rating, the authors concluding that the deluded participants 'were not concealing anything'. High scores were found on the DAS which did not differ between groups with persecutory and non-persecutory delusions. Authors consider that high DAS may aggravate and maintain any psychiatric condition.
8. Kinderman *et al*., 1992	1. Persecutory delusions (delusional disorder and schizophrenia) (23) 2. Depressed (depression and anorexia nervosa) (21) 3. Non-clinical control (28)	ASQ re-rating	N/A	N/A	√?	Study pooled data from previous Kaney & Bentall (1989) study and newly collected data. Independent judges re-rated causes previously generated by participants for internality. While deluded participants *only* rated own generated causes as external for negative events, ragings by judges were more internal. Authors conclude that deluded group has a particular tendency to bias their ratings of their own explanations of negative events towards the external, in order to defend self-esteem. Results thought to represent a form of self-deception, preventing deluded participants becoming aware of negatively self-referent attitudes (see text for discussion).
9. Peters *et al*., 1997	1. Deluded mixed diagnoses (23) 2. Depressed (22) 3. Non-clinical control (36)	PIT	N/A	√	N/A	This study investigated a range of cognitive processes in people with delusions, self-schema, cognitive inhibition and probabilistic reasoning. The PIT was considered to access self-schemata. On the PIT the deluded group did not differ from the control group, but did differ from depressed group, who showed expected internal attributions for negative events. Underlying low self-esteem in the deluded group could not be confirmed.

Study	Groups of participants (*N*)	Measure/task	Results: Overt normal/high self-esteem	Results: Covert low self-esteem	Results: Discrepancy	Comments
10. Bentall & Kaney, 1996	1. Depressed persecutory delusions (delusional disorder) (10) 2. Non-depressed persecutory delusions (delusional disorder) (10) 3. Depressed (20) 4. Non-clinical (20)	1. DAS 2. SRIRT	?	N/A	× ?	The groups with depression and delusions had higher DAS scores than non-clinical controls, the highest scores in the depressed group. Authors argue that high DAS scores imply self-discrepancies. On the SRIRT both deluded groups and non-clinical control groups endorsed significantly more positive than negative items than the depressed group and authors argue this represents a positive self-presentation bias. However, although analysis of this is not reported, visual inspection of the data suggests that both deluded groups endorsed numerically more negative words than normal controls and fewer than depressed group. The positive self-presentation is more marked in the non-clinical control group. Overall, the recall data showed no bias in any of the groups. An analysis of endorsed words only showed that normal controls differed from other groups in recalling more positive than negative words. The authors acknowledge that the SRIRT data are equivocal.

ASQ = Attributional Style Questionnaire; CSEI = Coopersmith Self Esteem Inventory; DAS = Dysfunctional Attitudes Scale; EST = Emotional Stroop Task; PIT = Pragmatic Inference Task; PPQ = Personal Profile Questionnaire; PQ = Personal Qualities Questionnaire (Higgins); RSEQ = Rosenberg Self Esteem Questionnaire; SCQ = Self Concept Questionnaire (Robson); SRIRT = Self Referent Incidental Recall Task.

Esteem Inventory (Coopersmith, 1984; used by Candido & Romney, 1990), the Rosenberg Self-Esteem Questionnaire (Rosenberg, 1965; employed by Lyon *et al*., 1994) and the Self-Concept Questionnaire (Robson, 1989; used by Freeman *et al*., 1998). These questionnaires set out to measure global self-esteem (which is hypothesised to derive from self-discrepancies) rather than measuring a number of specific domains of the self. However, the Robson questionnaire does measure a number of domains. Secondly, Kinderman & Bentall (1996*a*) have used an adaptation of Higgins *et al*.'s (1986) Selves Questionnaire, the Personal Qualities Questionnaire, which involves generating personal attributes and which assesses actual-self, ideal-self and ought-self perceptions and also considers the standpoint of others. Finally, two studies which examined discrepancies (Bentall & Kaney, 1996; Kinderman, 1994) used methods involving the endorsement of positive and negative adjectives as self-descriptive, which provide an informal view of the self-concept; in both studies, the endorsement phase was a precursor to a second phase investigating covert processes.

There are seven studies which, in various ways, have examined covert self-esteem, five of which have also investigated whether there is a discrepancy between overt and covert self-esteem (see also Table 5). It is a methodological challenge to elicit covert self-evaluations. The theory under discussion proposes that the individual is motivated to prevent these cognitions reaching consciousness and thus studies must employ methods of penetrating the defence, accessing cognitions of which the participant is neither aware nor wishes to be aware.

First, three studies have used the Emotional Stroop Task (Bentall & Kaney, 1989; Fear *et al*., 1996; Kinderman, 1994). The Emotional Stroop Task is thought to measure preconscious or automatic cognitive processes and has been widely used as a measure of covert attentional bias, signalling the emotional salience of words by slowed naming of the colours in which salient, as opposed to neutral or non-salient, words are printed (Brewin, 1988). The studies described here have used as stimuli strings of Os, neutral words and, variously, depression-related, threat-related, anxiety-related and positive and negative self-descriptive words. Studies have shown the Stroop effect even when participants are not consciously aware of the material presented (Williams, Mathews & MacLeod, 1996) and it can therefore be regarded as a valid measure of individuals' covert concerns.

The second approach to attempting to 'penetrate the defence' was taken by Lyon *et al*. (1994), drawing on earlier work with people with mania (Winters & Neale, 1985). Peters *et al*. (1997) have attempted to replicate this work. In Lyon *et al*.'s original study, responses on a parallel form of the ASQ (ASQpf) are compared with responses to the Pragmatic Inference Task (PIT), a questionnaire disguised as a memory task but based on the original ASQ, and requiring respondents to make attributions. The ASQpf served as a measure of explicit attributions while the PIT served as a measure of implicit attributions. This seems to the present authors a valid approach, although it should be noted that it is attributions rather than self-evaluations which are tapped.

The third approach to assessing covert processes involves studies of ratings of internality on the ASQ (Fear *et al*., 1996; Kinderman, Kaney, Morley & Bentall, 1992) Kinderman *et al*. (1992) reanalysed the responses on the ASQ made by

participants with persecutory delusions. Independent judges rated the causes, previously generated by the deluded participants, for internality or externality. (In the ASQ, participants first generate a cause and then rate it on the three dimensions.) Kinderman *et al*. (1992) argue that if judges rate the causes generated for negative events as more internal than the deluded participants' own ratings, this may reflect a defensive process, protecting the deluded person from negatively self-referent attributions. A discrepancy between the actual and ideal self-concept is inferred. This seems to the present authors to be an indirect and rather weaker test of the hypothesis of a defended negative self-concept, since an explanation of biases in internality ratings requires a further hypothesis. In Kinderman *et al*.'s case this is that the participants' first response (stating a cause of an event) is answered non-defensively, whereas the subsequent rating for internality activates the defence. This process is repeated many times on the ASQ; it seems just as plausible that the defensive responding would also occur in response to stating the cause, at least after a few responses. A more parsimonious explanation of the results might be that the internality ratings simply reflect an overt externalising bias. The unreliability of the internality dimension is also noteworthy in this context (Reivich, 1995). Whereas defensive processes may be implicated, it is possible that the deluded participants differed for other non-defensive reasons in their use of the rating scale.

Finally, Bentall & Kaney (1996) used two 'indirect assessments of self-schemata' selected to access negative self-representations even with highly defensive individuals. The measures used were the Dysfunctional Attitudes Scale (DAS; Weissman & Beck, 1978 (also employed by Fear *et al*., 1996)) and an incidental recall task, the Self-Referent Incidental Recall Task (SRIRT). Both measures have previously been used to study the role of self-schemata in depression. The DAS measures beliefs or attitudes which delineate excessively rigid and perfectionistic criteria for evaluating personal performance and self-worth. Bentall & Kaney (1996) propose that attitudes of this sort are likely to make one vulnerable to experiencing discrepancies between self-representations and self-ideals. However, it is not clear that the DAS is an indirect measure of self-schemata (i.e. that it assesses implicit/covert self-esteem), as is claimed: as a self-report measure, it attempts to assess self-schemata directly. It is also not clear that vulnerability to self-discrepancies implies active self-discrepancies. A direct measure of self-representation employed was a list of positive (success-related) and negative (failure-related) adjectives, which participants were asked to endorse as self-descriptive. A surprise recall task (SRIRT) of these lists was then used to access self-schemata covertly. The present authors consider that the SRIRT is likely to be a valid method of accessing covert self-evaluations, since it is based on evidence of the self-referent encoding effect, the general tendency for superior recall of words which have been encoded in relation to self-schemata (see Bentall & Kaney, 1996).

Results from the studies of discrepancies in overt and covert self-esteem

Table 5 first shows the direct studies of overt self-esteem. The findings are mixed from the four direct studies of overt self-esteem and persecutory delusions. Lyon *et al*. (1994) and Kinderman & Bentall (1996*a*) have found support for their proposal

of normal or high self-esteem/low self-discrepancies, whereas Freeman *et al*. (1998) found low self-esteem across all the Robson domains, both in people with persecutory delusions and in other patients with positive symptoms of psychosis. Candido & Romney (1990) found self-esteem to be normal in a non-depressed paranoid group, but to be low in a depressed paranoid group. Self-esteem therefore differed with levels of depression, as it did in the Lyon *et al*. study. Freeman *et al*. (1998) examined the relationship with depression. In people with persecutory delusions, self-esteem and depression were negatively correlated, cross-sectionally and longitudinally, which was considered to be consistent with normal emotional processes rather than defensive processes. However, they found a subgroup (30%) of people with persecutory delusions had normal levels of self-esteem. The two other studies, which used a more informal measure of the endorsement of positive and negative words as self-descriptive, had equivocal findings (Bentall & Kaney, 1996; Kinderman, 1994). Taking all six studies together the evidence does not strongly support the hypothesis of high or normal self-esteem in people with persecutory delusions, but may suggest that there are subgroups of people with persecutory delusions with normal self-esteem and others with low self-esteem.

Of course, finding evidence of low self-esteem in people with persecutory delusions does not preclude the weaker version of the delusion-as-defence account in which the delusion prevents self-esteem falling further. A test of this rests more squarely on evidence of a discrepancy between overt and covert self-esteem. However, with the 'weaker' version, one would expect a correlation over time between levels of self-esteem and the degree of conviction in the persecutory delusion, so that if the conviction diminishes, self-esteem is also lowered. Freeman *et al*. (1998) examined such data longitudinally and found little evidence for the predicted association, taking this as evidence against both the 'strong' and the 'weak' versions of the delusion-as-defence account.

Six studies have attempted to penetrate the defence and examine covert self-esteem. Only two of these studies have found convincing evidence of an implicit low self-esteem (Kinderman, 1994; Lyon *et al*., 1994), using the Emotional Stroop Task and the Pragmatic Inference Task respectively. Other studies employing these tasks have not confirmed the presence of covert negative self-esteem. Peters *et al*. (1997), in a study with a group of people with a variety of types of delusions, found no covert internalising bias on the PIT for negative events. On the EST, both Bentall & Kaney (1989) and Fear *et al*. (1996) found an attentional bias to threat and not to depressive words in the persecutory deluded group, findings which are not consistent with covert low self-esteem. Finally, the data from the Bentall & Kaney (1996) self-referent recall task did not find clear evidence of biases in recall reflecting a covert negative self-concept.

However, finding evidence of a covert negative self-concept is not sufficient to support the proposal of persecutory delusions as defence. This requires evidence of an overt–covert discrepancy. Five studies have attempted to examine this (see Table 5). Only one study, that by Lyon *et al*. (1994) of explicit and implicit attributions, provides clear evidence of a discrepancy with a method which is of high validity (although of attributions rather than directly of self-concepts, as discussed above). The findings of all the other studies are less clear. Kinderman *et al*. (1992), in their

study of independent judges re-rating deluded participants' internality ratings, have findings which are intriguing but are difficult to interpret, as discussed above; furthermore they have not been replicated by Fear *et al*. (1996). The present authors do not consider high scores on the DAS (found by Fear *et al*., 1996 and Bentall & Kaney, 1996) to provide convincing evidence of active covert self-discrepancies, as discussed above. Fear *et al*. (1996) suggest that highly dysfunctional attitudes might be found in a variety of psychiatric conditions, citing evidence of a high DAS score in people with obsessive compulsive disorder. Other researchers have found DAS scores to be raised in symptomatic and remitted patients with a variety of psychiatric disorders (e.g. Hill *et al*., 1989; Silverman *et al*., 1984). Fear *et al*. (1996) interpret the high DAS scores of their deluded participants as reflecting dysfunctional attitudes which might be expected to maintain any psychiatric condition. However, Bentall & Kaney (1996) do consider DAS scores as reflecting self-discrepancies. Finally, Kinderman (1994) interprets his findings as evidence of discrepancies. The present authors differ in their interpretation of his study, and are doubtful that discrepancies are shown.

Kinderman (1994) compares the results of an overt measure (endorsement of positive and negative adjectives as self-descriptive) with the covert measure (the EST of attentional bias to positive and negative adjectives). The present authors accept Kinderman's conclusion that there is evidence of a covert negative self-concept from the results of the EST in the persecutory group; this may, however, be accounted for by the levels of depression in the group. The interpretation of the pattern of endorsement of adjectives shown by the persecutory group in the overt phase is more complex. Kinderman analysed two sets of data, the endorsement as self-descriptive of the words from the whole questionnaire and the endorsement of a smaller group of words subsequently used in the EST. Whereas the deluded group endorsed similar numbers of positive words as the control group, they endorsed similar numbers of negative words as the depressed group; this latter finding is clearer when all words are considered, where the persecutory group and the depressed group do not differ from each other but differ from controls, while for the target words the persecutory group fell between the control and depressed group, differing from neither. Therefore, on overt self-rating, the deluded group did not differ from a depressed group in terms of the extent to which low self-esteem words were seen as self-descriptive; consequently, the interference of such words in the EST does not seem to the present authors to be discrepant and to provide evidence for a defence. Kinderman, however, concludes that a discrepancy has been found. He differs in his interpretation by emphasising, not the actual endorsement of low self-esteem words, but the relative rate of endorsement of positive and negative words: the deluded group rated more positive words as self-descriptive than the depressed group. He therefore concludes that the deluded participants' rating of the negative words is 'significantly less self-descriptive' (p. 63) than the depressed participants' ratings. This does not seem to the present authors to follow: if negative words are endorsed, they will be salient regardless of the endorsement of positive words. Indeed, the slowing of the deluded group on the EST by both positive and negative words is entirely consistent with this.

Most of the studies reviewed here have employed groups of people with

persecutory delusions. Are the findings specific to such delusions? Only Fear *et al*. (1996) and Freeman *et al*. (1998) compared groups of people with psychosis with and without persecutory delusions. Freeman *et al*. (1988) found no differences between their groups on measures of overt self-esteem, and Fear *et al*. (1996) found few differences on a range of measures between their deluded groups. There is no evidence to indicate that the processes investigated here are specific to persecutory delusions. Similarly, the present authors cannot draw any conclusions from these studies about the role of diagnosis. Most studies have employed diagnostically mixed groups and have not systematically investigated differences by diagnosis. Finally, whether IQ or gender differences affect these findings has not been systematically studied, although most studies have attempted to match groups by gender.

Do people with persecutory delusions have positive self-esteem and self-discrepancies? Do persecutory delusions serve a defence?

A very impressive series of studies is discussed above, largely conducted by Bentall, Kaney and Kinderman and colleagues, but supplemented by the work of other groups. Overall the findings of these 10 studies are variable and a consistent picture does not emerge, whether of levels of overt self-esteem or of self-discrepancies. First, in terms of the data on overt self-esteem, it is important to consider the role of depression. In all three studies which examined this, self-esteem scores were inversely related to depression scores (Candido & Romney, 1990; Freeman *et al*., 1998; Lyon *et al*., 1994). This is what would be expected in the general population and does not reflect the 'paradoxical combination of high self-esteem and high depression' proposed by Bentall (1994). Freeman *et al*. (1998) have argued that where depression and self-esteem scores are related in this way, this favours an account which invokes the role of 'normal emotional processes', in which self-esteem and depression are closely and reciprocally related, rather than an account based on defensive processes.

Secondly, the studies here have not established that overt–covert self-esteem discrepancies are present, which is the central prediction of both the 'strong' and the 'weak' versions of the Bentall and colleagues' theory. To demonstrate such a proposal is challenging. The present authors have proposed that the methods in some of the studies reviewed are of questionable validity and therefore the interpretation of the results is problematic. The review of these studies finds support for overt–covert self-esteem discrepancies in one study, but in four further studies the authors consider the balance of evidence is against discrepancies in two and uncertain in two. In view of this, they cannot conclude that there is yet strong empirical support for the persecutory delusions as a defence theory. An externalising attributional bias for negative events has been convincingly demonstrated in people with persecutory delusions. This element of the theory is well supported by evidence. However, an externalising bias for negative events does not necessarily serve a defensive function. The present critical analysis of a number of studies suggests that although a proportion of people with persecutory delusions may have normal or high self-esteem, this proportion may be less than half. Furthermore they are often depressed, which is not predicted by the re-worked theory of Bentall *et al*. (1994),

since dejection-related emotions are not expected if actual–ideal self-discrepancies are minimised by a defence. These findings of low self-esteem and depression are also important clinically. The evidence does not confirm crucially that overt–covert self-esteem discrepancies are present in the majority of cases; however, the present authors do consider that the equivocal results reviewed in the preceding section argue for the possibility that the defence account applies to a subgroup.

CONCLUSIONS

A number of empirical studies investigating three different accounts of delusion have been reviewed above. None of the accounts, as originally formulated, is fully supported. However, considerable progress has been made in recent years in replicating experimental results which provide evidence of cognitive processes implicated in delusions. First, the present authors have found strong support for a reasoning bias in people with delusions which is best described as a data-gathering bias, a tendency for people with delusions to gather less evidence than controls so that they jump to conclusions. This may be exaggerated with emotive or self-referent material but is also present with material which does not apparently relate to the individual's current concerns. This does not lead directly to errors, but may facilitate the early acceptance of incorrect hypotheses. Secondly, there is strong evidence of an attributional bias in people with persecutory delusions, which leads to externalising blame for negative events; there are early indications that this may result from a particular tendency to personalise—that is, to blame people rather than situations when things go wrong. However, whether this attributional bias functions in general to defend against underlying low self-esteem is more doubtful; the present authors do, however, consider that the defence account may apply to a subgroup of people with persecutory delusions. It is also not established whether these attributional biases are specific to persecutory as opposed to other types of delusions. Finally, a number of recent studies suggests that people with persecutory and other delusions may be poor at representing the mental states of others, although this deficit may be related to a more general reasoning factor.

There are a number of general methodological and conceptual issues which arise from this review. The authors note that few studies provide details on the process of participant recruitment; for example refusals to participate are very rarely reported. The studies also differ a great deal in the characteristics of participants recruited. For example, studies may recruit in-patients, or males, or groups with delusions and low depression scores. Many of the studies do not systematically investigate the possible effects of such variables on the results. Whereas Bentall and colleagues have paid attention to the role of depression, and Frith and colleagues to IQ, other variables which may plausibly contribute to the data such as gender, ethnicity, medication, overall psychopathology, anxiety and the duration of illness have been less systematically studied. The reliability of methods for allocating participants to groups on the basis of delusional subtypes is uncertain. Many studies do not report explicit criteria for group allocation. In some studies the terms 'paranoid' and 'persecutory' are used interchangeably and without definition, and there is no general agreement about the grouping of different subtypes of delusion

(so that studies may combine in a 'persecutory' group delusions of persecution and reference, persecutory and grandiose delusions or persecutory and passivity delusions). The studies also often omit to clarify whether participants hold concurrently a number of different types of delusion. The application of clear criteria for the membership of the category of persecutory delusion is needed.

A key conceptual issue arising from this review concerns the specificity of the different accounts. Frith's approach is the most specific. His theory of mind proposals are addressed at delusions as symptoms of schizophrenia and at specific subtypes of delusion: persecution and reference, but not passivity. However, he does allow that the same deficits may also be found in patients with negative symptoms and not delusions. Bentall and colleagues have specified that their account applies to people with persecutory delusions, regardless of diagnosis, but have not set out to examine whether their findings are specific to persecutory delusions (although Fear and colleagues have investigated this). Garety and colleagues' account is the least specific: they have made no proposals concerning delusional subtypes and in an investigation of diagnostic specificity found no differences in reasoning between participants with a diagnosis of schizophrenia and of delusional disorder. They also explicitly invoke a variety of cognitive processes in a multifactorial account.

The approaches reviewed can be seen to overlap in the phenomena they purport to explain. The accounts are also not contradictory. It is therefore quite possible that the different processes highlighted by the different theories may co-occur and may even interact. For example, Bentall & Kinderman (1998) have speculated that theory of mind deficits may be related to personalising attributional biases. Similarly, data-gathering biases may plausibly interact with a biased attributional style to result in hasty decisions about the intentions of others. Questions also remain to be answered about the applicability of these findings of cognitive biases and deficits to other symptoms of psychosis (addressed by Frith and colleagues and to a lesser extent by Garety and colleagues), to non-persecutory delusions and to delusions in specific diagnostic contexts. That perceptual processes are probably a central factor in some delusions, as Maher (1974) argued, is also important. The present authors consider a multifactorial account is consistent with the findings to date. However, future research should examine these different processes together and set out to specify with more precision which processes are likely to be active in which people with delusions.

Most of the studies reviewed are cross-sectional and demonstrate associations rather than cause. The processes examined may serve as maintaining factors or even consequences of current delusional ideation rather than as causes. Approaches which study patients in remission (e.g. Corcoran *et al*., 1995), which use longitudinal methodology (e.g. Freeman *et al*., 1998), or non-clinical participants with high delusional ideation (e.g. Linney *et al*., 1998) are valuable in mapping the relationship to these cognitive processes over time. Interestingly, Frith and colleagues have thereby convincingly demonstrated that theory of mind deficits are not present in patients in remission, which suggests that these deficits are not trait variables which predispose people to develop delusions.

It is noted in the introduction above that the single symptom approach has been commended as offering promise for the future. How has it fared in this review?

Persons (1986) listed a number of advantages of this approach: the avoidance of the misclassification of participants; the study of important phenomena which are often ignored; the facilitation of theoretical development; the isolation of single elements of pathology for study; the recognition of the continuity of clinical phenomena with normal phenomena; and improvements in diagnostic classification. Apart from the first and last of these (avoidance of misclassification and improved diagnosis), it is apparent that the research reviewed here has demonstrated these benefits. However, it should be noted that Frith and colleagues have taken a *symptom* approach rather than a *single symptom* one. They have therefore been able not only to isolate single elements of pathology for study but also to contrast different symptoms. A single symptom approach has the disadvantages of ignoring the complexity of presentation of individuals with delusions: frequently clusters of symptoms co-occur (as do subtypes of delusions), and it is proposed here that future theoretical development should incorporate attention to both single symptoms and clusters of symptoms. Single case studies, which incorporate longitudinal methodology, also provide valuable data, as demonstrated by the case studies in Halligan & Marshall (1996).

An additional conceptual issue concerns the multidimensional nature of delusions. This view proposes that delusions have a number of important characteristics, such as conviction, distress, preoccupation and action, which are relatively independent of each other and are found to differ along continua (Garety & Hemsley, 1994; Strauss, 1969). The research reviewed here mostly takes a categorical view, treating delusions as present/absent and, in line with standard diagnostic approaches, employing only the characteristic of conviction. The fluctuations in different dimensions are ignored. Cognitive accounts of delusion which consider other characteristics, such as the clinically important one of distress, may have much to offer. In their studies of hallucinations Chadwick & Birchwood (1994) and Close & Garety (1998) have examined how appraisals of the content and meaning of hallucinations predict distress. The present authors have recently taken this approach with delusional distress (Freeman & Garety, in press). They suggest that a fuller account of delusions which can give due credit to the multidimensional nature of delusions will incorporate content as well as cognitive processes, for example by considering the individual's appraisals of their delusional beliefs.

Some implications can be drawn for cognitive therapy approaches from this work. The studies of a variety of reasoning processes, which found that people with delusions are competent at testing hypotheses and probabilistic reasoning, are important in that these abilities can be used in cognitive therapy. On the basis of the evidence reviewed, therapy for delusions should address any tendency to jump to conclusions, aim to work on personalising or externalising attributions and enhance social understanding. The mixed findings on depression and self-esteem suggest they should be assessed carefully and interventions tailored accordingly. This is in keeping with Trower & Chadwick's (1995) recommendations. They suggest that, for patients with explicit low self-esteem and high levels of depression, the intervention should focus on negative self-evaluative beliefs before addressing delusional beliefs, an approach which is not relevant where negative self-evaluations are not apparent.

This review has also highlighted the need for further theoretical and empirical development of the understanding of the possible direct rather than defensive

relationship between emotional processes, such as depression and anxiety, and delusions. Consistent with this approach, Birchwood & Iqbal (1998) have argued that depression and loss of self-esteem in some people with psychosis may be viewed as a reaction to the experience of psychosis as an uncontrollable traumatic life event. The need for further work on this is apparent from a recent study of cognitive-behavioural therapy for psychosis (Kuipers *et al*., 1998) which found improvements in delusions and hallucinations, and reductions in delusional distress, but no changes in levels of depression or self-esteem.

What are the important future directions of this research into cognitive processes in delusions? The authors consider the major outstanding issues to be clarifying the aetiological status of findings, examining the co-occurrence of the different cognitive processes, clarifying the specificity of the theories, and investigating more fully the contribution of a number of variables such as IQ, medication and depression to the effects found. They suggest that research which uses patients with early psychosis, non-clinical samples with high delusional ideation or longitudinal methods will be informative. Investigations which bring together attributions, theory of mind tasks and non-social reasoning will clarify their relationships. A multidimensional approach to delusions which examines content, cognitive processes and emotional processes is advocated.

The evidence reviewed here is consistent with a multifactorial account of delusions. It is probable that delusions, understood as attempts to explain experiences or events, develop against a background of a person's existing personality and beliefs and as a result of a combination of alterations or biases in perception, affect and judgment. A multifactorial account may be the best available, but it is not informative unless more precise relationships can be specified and predictions made. It will be with Marshall & Halligan's (1996) 'healthy discord' of controversy and disagreement that one may hope to arrive at some better understanding.

REFERENCES

Abroms, G. M., Taintor, Z. C. & Lhamon, W. T. (1966). Percept assimilation and paranoid severity. *Archives of General Psychiatry*, **14**, 491–496.

Alloy, L. B. & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, **91**, 112–149.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders*, 3rd ed., revised. Washington, DC: American Psychiatric Association.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders*, 4th ed. Washington, DC: American Psychiatric Association.

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

Bentall, R. P. (1994). Cognitive biases and abnormal beliefs: Towards a model of persecutory delusions. In A. S. David & J. Cutting (Eds), *The neuropsychology of schizophrenia*, pp. 337–360. London: Lawrence Erlbaum.

Bentall, R. P. & Kaney, S. (1989). Content specific processing and persecutory delusions: An investigation using the emotional Stroop test. *British Journal of Medical Psychology*, **62**, 355–364.

Bentall, R. P. & Kaney, S. (1996). Abnormalities of self-representation and persecutory delusions: A test of a cognitive model of paranoia. *Psychological Medicine*, **26**, 1231–1237.

Bentall, R. P., Kaney, S. & Dewey, M. E. (1991). Persecutory delusions: An attribution theory analysis. *British Journal of Clinical Psychology*, **30**, 13–23.

Bentall, R. P. & Kinderman, P. (1998). Psychological processes and delusional beliefs: Implications for the treatment of paranoid states. In T. Wykes, N. Tarrier & S. Lewis (Eds), *Outcome and innovation in psychological treatment of schizophrenia*, pp. 119–144. Chichester: Wiley.

Bentall, R. P., Kinderman, P. & Kaney, S. (1994). The self, attributional processes and abnormal beliefs: Towards a model of persecutory delusions. *Behaviour Research and Therapy*, **32**, 331–341.

Bentall, R. P. & Young, H. F. (1996). Sensible hypothesis testing in deluded, depressed and normal subjects. *British Journal of Psychiatry*, **168**, 372–375.

Birchwood, M. & Iqbal, Z. (1998). Depression and suicidal thinking in psychosis: A cognitive approach. In T. Wykes, N. Tarrier & S. Lewis (Eds), *Outcome and innovation in psychological treatment of schizophrenia*, pp. 81–100. Chichester: Wiley.

Birchwood, M., Mason, R., MacMillan, F. & Healy, J. (1993). Depression, demoralization and control over psychotic illness: A comparison of depressed and non-depressed patients with a chronic psychosis. *Psychological Medicine*, **23**, 387–395.

Bourchard, S., Vallières, A., Roy, M. & Maziade, M. (1996). Cognitive restructing in the treatment of psychotic symptoms in schizophrenia: A critical analysis. *Behavior Therapy*, **27**, 257–277.

Brett-Jones, J., Garety, P. & Hemsley, D. (1987). Measuring delusional experiences: A method and its application. *British Journal of Clinical Psychology*, **26**, 257–265.

Brewin, C. R. (1988). *Cognitive foundations of clinical psychology*. Hove: Lawrence Erlbaum.

Byrne, A. & MacLeod, A. K. (1997). Attributions and accessibility of explanations for future events in anxiety and depression. *British Journal of Clinical Psychology*, **36**, 505–520.

Cahill, C. & Frith, C. (1996). False perceptions or false beliefs? Hallucinations and delusions in schizophrenia. In P. W. Halligan & J. C. Marshall (Eds), *Method in madness*, pp. 267–291. Hove: Psychology Press.

Candido, C. L. & Romney, D. M. (1990) Attributional style in paranoid versus depressed patients. *British Journal of Medical Psychology*, **63**, 355–363.

Chadwick, P. D. J. & Birchwood, M. J. (1994). The omnipotence of voices: A cognitive approach to hallucinations. *British Journal of Psychiatry*, **164**, 190–201.

Chadwick, P. D. J. & Lowe, C. F. (1990). Measurement and modification of delusional beliefs. *Journal of Consulting and Clinical Psychology*, **58**, 225–232.

Chapman, L. J. & Chapman, J. P. (1988). The genesis of delusions. In T. F. Oltmanns & B. A. Maher (Eds), *Delusional beliefs*, pp. 167–183. New York: Wiley.

Close, H. & Garety, P. (1998). Cognitive assessment of voices: Further developments in understanding the emotional impact of voices. *British Journal of Clinical Psychology*, **37**, 173–188.

Colby, K. M., Faught, W. S. & Parkinson, R. C. (1979). Cognitive therapy of paranoid conditions: Heuristic suggestions based on a computer simulation. *Cognitive Therapy and Research*, **3**, 55–60.

Coopersmith, S. (1984). *Self-esteem inventories*. Palo Alto, CA: Consulting Psychologists Press.

Corcoran, R., Cahill, C. & Frith, C. D. (1997). The appreciation of visual jokes in people with schizophrenia: A study of 'mentalizing' ability. *Schizophrenia Research*, **24**, 319–327.

Corcoran, R., Mercer, G. & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: Investigating 'theory of mind' in people with schizophrenia. *Schizophrenia Research*, **17**, 5–13.

Doody, G. A., Gotz, M., Johnstone, E. C., Frith, C. D. & Cunningham Owens, D. G. (1998). Theory of mind and psychoses. *Psychological Medicine*, **28**, 397–405.

Dudley, R. E. J., John, C. H., Young, A. W. & Over, D. E. (1997*a*). Normal and abnormal reasoning in people with delusions. *British Journal of Clinical Psychology*, **36**, 243–258.

Dudley, R. E. J., John, C. H., Young, A W. & Over, D. E. (1997*b*). The effect of self-referent material on the reasoning of people with delusions. *British Journal of Clinical Psychology*, **36**, 575–584.

Edwards, W. (1982). Conservatism in human information processing. In D. Kahneman, P. Slovic & A. Tversky (Eds), *Judgement under uncertainty: Heuristics and biases*, pp. 359–369. Cambridge: Cambridge University Press.

Fear, C. F. & Healy, D. (1997). Probabilistic reasoning in obsessive-compulsive and delusional disorders. *Psychological Medicine*, **27**, 199–208.

Fear, C., Sharp, H. & Healy, D. (1996). Cognitive processes in delusional disorders. *British Journal of Psychiatry*, **168**, 1–8.

Fischhoff, B. & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, **90**, 239–260.

Freeman, D. & Garety, P. (in press). Worry, worry processes and dimensions of delusions: An exploratory investigation of a role for anxiety processes in the maintenance of delusional distress. *Behavioural and Cognitive Psychotherapy*.

Freeman, D., Garety, P., Fowler, D., Kuipers, E., Dunn, G., Bebbington, P. & Hadley, C. (1998). The London–East Anglia randomised controlled trial of cognitive-behaviour therapy for psychosis IV: Self esteem and persecutory delusions. *British Journal of Clinical Psychology*, **37**, 415–430.

Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine*, **17**, 631–648.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove: Lawrence Erlbaum.

Frith, C. D. & Corcoran, R. (1996). Exploring 'theory of mind' in people with schizophrenia. *Psychological Medicine*, **26**, 521–530.

Frith, C. D. & Done, D. J. (1989). Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychological Medicine*, **19**, 359–363.

Garety, P. A. (1991). Reasoning and delusions. *British Journal of Psychiatry*, **159** (Suppl. 14), 14–18.

Garety, P. A. (1992). Making sense of delusions. *Psychiatry*, **55**, 282–291.

Garety, P. A., Fowler, D., Kuipers, E., Freeman, D., Dunn, G., Bebbington, P. E., Hadley, C. & Jones, S. (1997). The London–East Anglia randomised controlled trial of cognitive-behaviour therapy for psychosis II: Predictors of outcome. *British Journal of Psychiatry*, **171**, 420–426.

Garety, P. A. & Hemsley, D. R. (1987). Characteristics of delusional experience. *European Archives of Psychiatry and Neurological Sciences*, **236**, 294–298.

Garety, P. A. & Hemsley, D. R. (1994). *Delusions: Investigations into the psychology of delusional reasoning.* Hove: Psychology Press.

Garety, P. A., Hemsley, D. R. & 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.

Halligan, P. W. & Marshall, J. C. (Eds) (1996). *Method in madness*. Hove: Psychology Press.

Hemsley, D. R. (1987). An experimental psychological model for schizophrenia. In H. Hafner, W. F. Gattaz & W. Janzarik (Eds), *Search for the causes of schizophrenia*, pp. 179–188. Heidelberg: Springer.

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

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, **94**, 319–340.

Higgins, E. T., Bond, R., Klein, R. & Strauman, T. J. (1986). Self discrepancies and emotional vulnerability: How magnitude, accessibility and type of discrepancy influence affect. *Journal of Personality and Social Psychology*, **41**, 1–15.

Hill, C. V., Oei, Tian, P. & Hill, M. A. (1989). An empirical investigation of the specificity and sensitivity of the Automatic Thoughts Questionnaire and Dysfunctional Attitudes Scale. *Journal of Psychopathology and Behavioural Assessment*, **11**, 291–311.

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

Jaspers, K. (1913/1963). *General psychopathology* (trans. J. Hoenig & M. Hamilton). Manchester: Manchester University Press.

John, C. H. & Dodgson, G. (1994). Inductive reasoning in delusional thinking. *Journal of Mental Health*, **3**, 31–49.

Kaney, S. & Bentall, R. P. (1989). Persecutory delusions and attributional style. *British Journal of Medical Psychology*, **62**, 191–198.

Kaney, S. & Bentall, R. P. (1992). Persecutory delusions and the self-serving bias: Evidence from a contingency judgement task. *Journal of Nervous and Mental Disease*, **180**, 773–780.

Kemp, R., Chua, S., McKenna, P. & David, A. (1997). Reasoning and delusions. *British Journal of Psychiatry*, **170**, 398–405.

Kinderman, P. (1994). Attentional bias, persecutory delusions and the self-concept. *British Journal of Medical Psychology*, **67**, 53–66.

Kinderman, P. & Bentall, R. P. (1996*a*). Self-discrepancies and persecutory delusions: Evidence for a model of paranoid ideation. *Journal of Abnormal Psychology*, **105**, 106–113.

Kinderman, P. & Bentall, R. P. (1996*b*). A new measure of causal locus: The internal, personal and situational attributions questionnaire. *Personality and Individual Differences*, **20**, 261–264.

Kinderman, P. & Bentall, R. P. (1997). Causal attributions in paranoia and depression: Internal, personal and situational attributions for negative events. *Journal of Abnormal Psychology*, **106**, 341–345.

Kinderman, P., Kaney, S., Morley, S. & Bentall, R. P. (1992). Paranoia and the defensive attributional style: Deluded and depressed patients' attributions about their own attributions. *British Journal of Medical Psychology*, **65**, 371–383.

Kuipers, E., Fowler, D., Garety, P., Chisholm, D., Freeman, D., Dunn, G., Bebbington, P. & Hadley, C. (1998). London–East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis III: Follow-up and economic evaluation at 18 months. *British Journal of Psychiatry*, **173**, 61–68.

Kuipers, E., Garety, P. A., Fowler, D., Dunn, G., Bebbington, P. E., Freeman, D. & Hadley, C. (1997). The London–East Anglia randomised controlled trial of cognitive-behaviour therapy for psychosis I: Effects of the treatment phase. *British Journal of Psychiatry*, **171**, 319–327.

Langdon, R., Michie, P. T., Ward, P. B., McConaghy, N., Catts, S. V. & Coltheart, M. (1997). Defective self and/or other mentalising in schizophrenia: A cognitive neuropsychological approach. *Cognitive Neuropsychiatry*, **2**, 167–193.

Linney, Y., Peters, E. & Ayton, P. (1998). Reasoning biases in delusion-prone individuals. *British Journal of Clinical Psychology*, **37**, 247–370.

Lyon, H. M., Kaney, S. & Bentall, R. P. (1994). The defensive function of persecutory delusions: Evidence from attribution tasks. *British Journal of Psychiatry*, **164**, 637–646.

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 & B. A. Maher (Eds), *Delusional beliefs*, pp. 15–33. New York: Wiley.

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

Maher, B. A. & Ross, J. S. (1984). Delusions. In H. E. Adams & P. Sutker (Eds), *Comprehensive handbook of psychopathology*, pp. 383–411. New York: Plenum.

Marshall, J. C. & Halligan, P. W. (1996). Towards a cognitive neuropsychiatry. In P. W. Halligan & J. C. Marshall (Eds), *Method in madness*, pp. 3–11. Hove: Psychology Press.

McCormick, D. J. & Broekema, V. J. (1978). Size estimation, perceptual recognition and cardiac response rate in acute paranoid and non-paranoid schizophrenics. *Journal of Abnormal Psychology*, **87**, 385–398.

McReynolds, P., Collins, B. & Acker, M. (1964). Delusional thinking and cognitive organisation in schizophrenia. *Journal of Abnormal and Social Psychology*, **69**, 210–212.

Mitchley, N. J., Barber, J., Gray, J. M., Brooks, D. N. & Livingstone, M. G. (1998). Comprehension of irony in schizophrenia. *Cognitive Neuropsychiatry*, **3**, 2, 127–138.

Mortimer, A. M., Bentham, P., McKay, A. P., Quemada, I., Clare, L., Eastwood, N. & McKenna, P. J. (1996). Delusions in schizophrenia: A phenomenological and psychological exploration. *Cognitive Neuropsychiatry*, **1**, 289–303.

Mullen, P. (1979). Phenomenology of disordered mental function. In P. Hill, R. Murray & G. Thorley (Eds), *Essentials of post-graduate psychiatry*, pp. 22–54. London: Academic Press.

Oltmanns, T. F. (1988). Approaches to the definition and study of delusions. In T. F. Oltmanns & B. A. Maher (Eds), *Delusional beliefs*, pp. 3–12. New York: Wiley.

Oltmanns, T. F. & Maher, B. A. (Eds) (1988). *Delusional beliefs*. New York: Wiley.

Persons, J. B. (1986). The advantages of studying psychological phenomena rather than psychiatric diagnosis. *American Psychologist*, **41**, 1252–1260.

Peters, E., Day, S. & Garety, P. (1997). From preconscious to conscious processing: Where does the abnormality lie in delusions? *Schizophrenia Research*, **24**, 120.

Peterson, C., Semmel, A., Von Baeyer, C., Abramson, L., Metalsky, G. I. & Seligman, M. E. P. (1982). The Attributional Style Questionnaire. *Cognitive Therapy and Research*, **3**, 287–300.

Phillips, L. D. & Edwards, W. (1966). Conservatism in a simple probabilistic inference task. *Journal of Experimental Psychology*, **72**, 346–354.

Reivich, K. (1995). The measurement of explanatory style. In G. M. Buchanan & M. E. P. Seligman (Eds), *Explanatory style*, pp. 21–47. Hillsdale, NJ: Lawrence Erlbaum.

Robson, P. J. (1988). Self-esteem—a psychiatric view. *British Journal of Psychiatry*, **153**, 6–15.

Robson, P. J. (1989). Development of a new self-report questionnaire to measure self esteem. *Psychological Medicine*, **19**, 513–518.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Ross, L. & Anderson, C. A. (1982). Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessment. In K. Kahneman, P. Slovic & A. Tversky (Eds), *Judgement under uncertainty: Heuristics and biases*, pp. 129–152. New York: Cambridge University Press.

Salzinger, K. (1984). The immediacy hypothesis in a theory of schizophrenia. In W. D. Spaulding & J. K. Cole (Eds), *Theories of schizophrenia and psychosis. Nebraska symposium on motivation*, pp. 231–282. Lincoln: University of Nebraska Press.

Sarfati, Y., Hardy-Baylé, M. C., Besche, C. & Widlöcher, D. (1997). Attribution of intentions to others in people with schizophrenia: A non-verbal exploration with comic strips. *Schizophrenia Research*, **25**, 199–209.

Sharp, H. M., Fear, C. F. & Healy, D. (1997). Attributional style and delusions: An investigation based on delusional content. *European Psychiatry*, **12**, 1–7.

Silverman, J. S., Silverman, J. A. & Eardley, D. A. (1984). Do maladaptive attitudes cause depression. *Archives of General Psychiatry*, **41**, 28–30.

Slade, P. D. & Bentall, R. P. (1988). *Sensory deception: A scientific analysis of hallucination*. London: Croom Helm.

Strauss, J. S. (1969). Hallucinations and delusions as points on continua function. *Archives of General Psychiatry*, **20**, 581–586.

Trower, P. & Chadwick, P. (1995). Pathways to defense of the self: A theory of two types of paranoia. *Clinical Psychology: Science and Practice*, **2**, 263–278.

Weissman, A. N. & Beck, A. T. (1978). *Development and validation of the Dysfunctional Attitude Scale*. Paper presented at the Annual Meeting of the Association for the Advancement of Behavior Therapy, Chicago.

Williams, J. M. G., Mathews, A. & MacLeod, C. (1996). The Emotional Stroop Task and Psychopathology. *Psychological Bulletin*, **120**, 1, 3–24.

Wing, J. K., Cooper, J. E. & Sartorius, N. (1974). *The measurement and classification of psychiatric symptoms*. Cambridge: Cambridge University Press.

Winters, K. C. & Neale, J. M. (1983). Delusions and delusional thinking in psychotics: A review of the literature. *Clinical Psychology Review*, **3**, 227–253.

Winters, K. C. & Neale, J. M. (1985). Mania and low self-esteem. *Journal of Abnormal Psychology*, **94**, 252–290.

World Health Organisation (1992). *SCAN: Schedules for clinical assessment in neuropsychiatry*. Geneva: WHO.

Young, H. F. & Bentall, R. P. (1995). Hypothesis testing in patients with persecutory delusions: Comparison with depressed and normal subjects. *British Journal of Clinical Psychology*, **34**, 353–369.

Young, H. F. & Bentall, R. P. (1997*a*). Probabilistic reasoning in deluded, depressed and normal subjects: Effects of task difficulty and meaningful versus non-meaningful material. *Psychological Medicine*, **27**, 455–465.

Young, H. F. & Bentall, R. P. (1997*b*). Social reasoning in individuals with persecutory delusions: The effects of additional information on attributions for the observed behaviour of others. *British Journal of Clinical Psychology*, **36**, 569–573.

Zigler, E. & Glick, M. (1988). Is paranoid schizophrenia really camouflaged depression? *American Psychologist*, **43**, 284–290.

Zimbardo, P. G., Andersen, S. M. & Kabat, L. G. (1981). Induced hearing deficit generates experimental paranoia. *Science*, **212**, 1529–1531.

Received 17 March 1998; revised version received 27 November 1998