British Journal of Psychiatry (1991), **159** (suppl. 14), 36–41

Memory, Knowledge and Delusions

P. J. McKENNA

In schizophrenia, where they find their fullest expression, delusions are an extraordinary phenomenon. They can take a bewildering variety of forms, persecutory, grandiose, hypochondriacal, referential, and so on. They can be fixed and immutable, or alternatively may shift and change to the point of what Bleuler termed a delusional chaos. Finally, their degree of implausibility starts from the inherently unlikely, ranges through the patently absurd and bizarre, and not infrequently comes to violate common sense at its most elementary in fantastic delusions.

Knowledge of the phenomenology of delusions draws heavily from the classical accounts of Jaspers (1959) and Schneider (1949). According to Jaspers, 'delusions proper', that is delusions in schizophrenia, are essentially qualitatively different from normal beliefs: he considered that they were alien, ununderstandable, and the manifestation of a radical transformation in the basic processes of thinking. Jaspers also argued at length that schizophrenic delusions were primary, that is, they were not explicable in terms of any other pathological phenomenon like mood change or perceptual abnormality (see also Walker, 1991). Schneider was largely responsible for pointing out that some delusions (now referred to as delusions of reference, misinterpretation, misidentification and delusional perceptions) involved the attribution of erroneous significance to neutral events; both he and Jaspers, however, were clear that not all delusions could be understood as having such an attribute.

Attempts to explain the phenomenon of delusions have been sparse. Most accounts date from the classical era of psychiatry. As reviewed by Arthur (1964) and Winters & Neale (1983), these approaches invariably invoked the idea that delusions were secondary to some other change, e.g. in affect, in self-perception or in personality. Perhaps the worst example was the theory (usually attributed to Griesinger) that delusions were secondary to poor intellect and judgement; this was readily contradicted by the existence of highly intelligent deluded patients. The best example was perhaps Bleuler's (1911) suggestion that formal thought disorder, although not in itself completely explaining the development of delusions, could, by virtue of loosening of associations and illogicality, provide fertile ground for their elaboration. In fact, all the classical theories of delusions failed to account for their basic phenomenological attributes, particularly Jaspers' contention that they were a primary, psychologically irreducible phenomenon (see Walker, 1991).

The only contemporary theory of delusions is that of Maher (1974). He proposed that there exists a cognitive machinery for forming beliefs, which is brought into play whenever an individual is confronted with novel and inexplicable events. In such circumstances, a characteristic sequence of actions takes place: initial puzzlement gives way to an examination of the environment for related untoward events. This is followed by the generation of preliminary explanatory hypotheses which are examined in turn against the novel event until one shows signs of being capable of explaining it. In a flash of insight, this is then provisionally accepted as true, and the degree of certainty is increased by further confirmatory testing.

Maher argued that when an individual was faced with abnormal experiences, e.g. hallucinations, this cognitive machinery, although not itself defective, could only arrive at erroneous conclusions – an idea similar to the "garbage in, garbage out" adage in computer science. This approach clearly faces the same difficulty as those mentioned above, that delusions are made secondary to some other change, whereas phenomenologically they are primary. To some extent Maher circumvented this objection by proposing that when primary abnormal phenomena were absent, there might be a 'central neuropathology' which gave rise to a free-floating feeling of significance, which could then be elaborated into referential delusions. Maher's theory has another shortcoming in its prediction that delusions, as attempts to explain unprecedented events and experiences, should incline to the mundane and realistic. In practice, however, delusions seem very clearly to incline to the bizarre and fantastic.

Given the poverty and flaws of existing theories, new approaches to the explanation of delusions are required. What follows is a fuller version of one such attempt which was first outlined by McKenna (1987). It does not amount to a theory of delusions; nor can it be considered a formal model for delusion formation. It is perhaps best regarded as a possible avenue into the phenomenon of delusions, an attempt to provide them with what Berrios (1991) calls

https://doi.org/10.1192/S0007125000296451 Published online by Cambridge University Press

a neurophysiological address. The proposal is that delusions represent a dysfunction in the brain and cognitive systems which subserve memory, in particular long-term or secondary memory. The argument supporting such a view proceeds at two levels: first, that delusions are a consequence of dysfunction in the brain structures implicated in memory, and secondly that delusions represent a dysfunction of the cognitive systems involved in memory.

Delusions as a dysfunction of brain structures implicated in memory

Brain regions which appear, from the evidence of human lesions, to be involved in memory (i.e. long-term or secondary memory) include the hippocampus and its related structures like the mamillary bodies, the anterior thalamic nuclei and the prefrontal cortex (e.g. see Gray, 1982; Squire, 1987; Baddeley & Wilson, 1988). All of these are components of a well defined structural complex whose core is the hippocampal formation and which has come to be referred to as the septo-hippocampal system (Swanson, 1978).

The structure and function of the septo-hippocampal system

As reviewed by Gray (1982), the septo-hippocampal system comprises a complicated series of internal loops, which makes only surprisingly limited contact with other brain regions. The system has three main inputs, from secondary and tertiary sensory association cortex via the entorhinal cortex, from the prefrontal cortex, and from the amygdala. Its important outputs reach the ventral striatum and the hypothalamus; a third large efferent pathway projects via the mammillary bodies and anterior thalamic nuclei to the cingulate cortex – which then projects back to the septo-hippocampal system to form its most prominent internal loop system.

In addition, the septo-hippocampal system receives a rich innervation from the ascending monoamine systems, noradrenergic, serotonergic and dopaminergic pathways all reaching some or all of its key component structures (see Fig. 1).

A huge body of animal experimental data has made it clear that the septo-hippocampal system has no simple motor, sensory, learning or emotional function – complete removal of the hippocampus bilaterally in rats in fact produces no immediately evident deficits. Instead, it appears that in some general way the septo-hippocampal system promotes efficiency, flexibility and sophistication in learned behaviour. After septo-hippocampal system damage, animals (typically rats) show a complex set of behavioural changes in the direction of perseveration and disinhibition. They remain capable of learning and unlearning responses but this

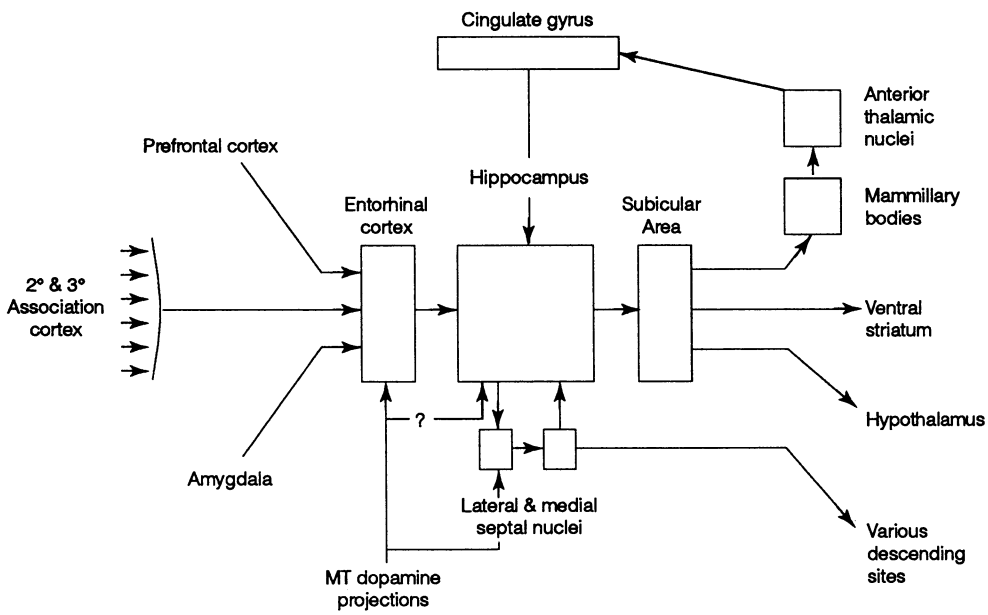


Fig. 1 Septo-hippocampal system connectivity. (M.T. Dopamine Projections = mesotelencephalic dopamine projections. NB: other monoamines project more widely to the system).

https://doi.org/10.1192/S0007125000296451 Published online by Cambridge University Press

becomes laborious, and insensitive to environmental change.

In a synthesis of the experimental findings and a reconciliation of their various interpretations, Gray (1982) proposed that the septo-hippocampal system acts quite generally to compare *actual* with *expected* in an animal's environment. This comparator function goes on continuously and enables behaviour to be facilitated or inhibited on an immediate basis. Via its sensory association area inputs, the system receives a continuous stream of *actual* sensory data. These are matched from moment to moment with *expected*, that is, predictions of what should be experienced in the circumstances based on previous experience, which are generated in the prefrontal and cingulate cortex. When *actual* and *expected* coincide, the system functions passively, permitting any current plan of motor responses to continue in its elaboration. At the same time, information that a 'match' has taken place is passed on for storage elsewhere (perhaps in the temporal lobe cortex). If, however, *actual* and *expected* fail to agree (e.g. because of a change in a previously stable environment), the septo-hippocampal system takes active control of behaviour. The currently engaged motor plan is inhibited, arousal abruptly increases, and alternative exploratory and investigative behaviours are instituted in an attempt to take in further information and resolve the difficulty that has interrupted responding.

Within the framework of Gray's scheme, monoamine inputs to the septo-hippocampal system serve to label incoming stimuli as 'important' in the sense of requiring particularly careful checking by the septo-hippocampal system. Specifically, noradrenalin inputs label stimuli as 'important, generally' (perhaps with particular respect to novelty) and serotonin inputs label stimuli as 'important, associated with punishment'. Both also enhance the mismatch detection function of the system, and bias its operation towards behavioural inhibition. The role of the dopamine inputs to the septo-hippocampal system was not specified by Gray (1982), but it can be argued that while they may also act to label incoming stimuli as 'important' (perhaps as 'important associated with reward') their effect would also be to antagonise behavioural inhibition and bias the system towards its passive match function.

In short, according to Gray, the septo-hippocampal system functions to check that events live up to expectations, and to intervene by inhibiting behaviour when they do not. The system also, in a less clearly specified way, participates in the storage and updating of information about the relationships between events and their consequences. Monoamine inputs to the system act essentially to draw attention to certain environmental events. In his initial formulation (Gray, 1982) the emphasis was on the consequences of behavioural inhibition which was argued to be an animal counterpart of the human experience of anxiety. In a subsequent elaboration (Gray *et al*, 1991) the emphasis has been shifted to the way in which the septo-hippocampal monitoring process interacts with information about the relationships of stimulus and reponse events ('stored regularities') and the running of motor programmes, so as to influence behaviour more generally.

Delusions as septo-hippocampal system dysfunction

A key part of Gray's (1982) theory is its application to human psychiatric disorders. Abnormal anxiety, he argues, is a consequence of excessive noradrenergic and/or serotonergic activity which results in activation and biasing of the septo-hippocampal system towards its mismatch function. More recently, taking a functional dopamine excess as a starting point, Gray *et al* (1991) attempted to explain some of the positive symptoms of schizophrenia as an alteration of the septo-hippocampal monitoring function on motor plans. The phenomenon of delusions, however, is scarcely touched on in this formulation.

A neglected aspect of both Gray's (1982) and Gray *et al*'s (1991) accounts is what the behavioural consequences might be if the septo-hippocampal system were, perhaps by virtue of excessive dopaminergic activity, to become abnormally biased towards its passive, match function. In these circumstances, it can be plausibly argued that the following would occur:

(a) Neutral stimuli would become erroneously identified as 'important'
(b) There would be a tendency for *expected* (i.e. predictions generated in the prefrontal cortex) to be erroneously judged as matching *actual* (i.e. sensory input concerning the current state of the environment) when they did not in fact coincide
(c) The erroneous information that a match had taken place would be passed on for storage and would alter the stored regularities used to make future predictions
(d) Plans of motor behaviour would continue to be elaborated, even though they were inappropriate for the environmental conditions.

What emerge from this line of speculation are abnormal phenomena with characteristics reminiscent of those of delusions. Point (a) states that events in the environment will acquire abnormal significance. Point (b) means essentially that predictions – or inductive inferences – will be erroneously verified as

https://doi.org/10.1192/S0007125000296451 Published online by Cambridge University Press

correct. Point (c) can be taken to indicate that knowledge concerning the relationship of events will be subverted, perhaps progressively so. Point (d) is more neutral, but if the term motor plans can be considered to apply to speech (as Gray and co-workers feel it can), the intriguing possibility is raised of abnormal speech acts being a consequence of septo-hippocampal disorder (see Berrios, 1991).

Delusions as a dysfunction of systems involved in memory

At first sight, there is no obvious link between delusions and memory, apart from their occasional conjunction in the phenomenon of delusional memory. Closer examination, however, reveals that there is a possible indirect link, in that both share some ground with the concept of knowledge. Belief can be equated with knowledge, to the extent that both refer to that which is known to be true. Memory can also be equated with knowledge, in the sense that both represent different forms of stored information. As made clear by Berrios (1991), to consider delusions as abnormal knowledge is certainly no more misleading than to consider them as abnormal beliefs. Perhaps more surprisingly, over the past two decades, the view that knowledge is closely related to memory has become not only legitimate, but also one of the most influential aspects of memory theory.

The concept of semantic memory

It is generally accepted that all information which needs to be held for more than a few seconds is held in a large capacity store termed long-term or secondary memory. In 1972, Tulving (see Tulving, 1983) proposed that long-term/secondary memory consisted of two conceptually distinct but closely interacting systems which he termed *episodic* and *semantic* memory. Episodic memory stores events from an individual's own experience which are personal and accompanied by a degree of specification as to time and place. Examples include what one had for breakfast, who one met on holiday last year, or what words were presented in a memory test. Semantic memory, in contrast, stores general knowledge about the world, concepts, and language; this information is impersonal and is held, as it were, in pure form, independent of any time and place specification. Examples here include knowing the capital of France, the chemical formula for salt, or, in the above-mentioned memory test, the meanings of the words themselves.

In some ways, episodic and semantic memory are similar: they are both concerned with the acquisition, retention and utilisation of information; they are both held outside consciousness in huge stores; and their contents are accessed, modified and updated by conscious mental activity in roughly similar ways. At the same time, a number of conceptual distinctions can be drawn, some of which are listed in Table 1. A further important difference is the apparent sparing of semantic memory in the amnesic syndromes seen in Korsakoff's psychosis or following bilateral medial temporal lobectomy. Such patients, while being completely unable to recall what they had for breakfast, any of the words in a list presented a few minutes earlier, etc., seem to have intact general knowledge and continue to know the meaning of words. Recently, however, a number of ingenious studies (e.g. reviewed by Squire, 1987) have suggested that there is an impairment in recently acquired semantic memory in the classical amnesic syndrome.

Delusions as a dysfunction of semantic memory

One point of view which can be advanced is that delusons are held in semantic memory. Since all concepts, beliefs and ideas are stored in semantic memory, this is easily defended, but is a theoretically empty, even tautological position. A stronger hypothesis is that delusions consist of an inappropriate laying down of new semantic memories. Semantic memory, like episodic memory, is clearly dynamic

Table 1
Differences between episodic and semantic memory (Modified from Tulving, 1983)

	Episodic memory	Semantic memory
Source of information	sensation	comprehension
Unit of information	event/episode	fact/idea/concept
Organisation of store	time and place	conceptual structure
Reference point	personal/self	impersonal/the world
Means of verification	intrinsic subjective veridicality	assessment of related facts
Capacity to infer	limited	rich
Recollective experience	remembered past	actualised knowledge
Retrieval report	remember	know

https://doi.org/10.1192/S0007125000296451 Published online by Cambridge University Press

in nature with new knowledge being acquired (and some old knowledge being lost) continually, in a process which must be governed in some way. If this process were to become deranged in such a way that its operation became unchecked, it could be envisaged that there would be a progressive accretion of erroneous semantic memories, false knowledge – or in other words delusions.

It is difficult to take this line of speculation further at a theoretical level. This is because the construct of semantic memory, despite its obvious importance, remains fragmentary and lacks a coherent conceptual framework (Baddeley, 1990). On the other hand, considerable progress in the experimental study of semantic memory has been made, and it is thus possible to examine the more general question of whether semantic memory is abnormal in schizophrenia. So far, such studies have been few and far between, but those that bear on the issue can be summarised as follows.

Koh and co-workers (see Koh, 1978) investigated various aspects of memory functioning in a sample of young acute schizophrenic patients. Semantic memory was examined by requiring the patients to sort cards showing different words into piles according to the categories to which they were considered to belong. Some of the words were obviously interrelated, whereas others were not related to any of the others. The schizophrenics were found to use fewer categories than controls, but the overall pattern of categorisation was not generally different between the two groups. Koh's conclusion was that semantic memory was unaffected in schizophrenia, but this was weakened by the use of inadequate diagnostic criteria and the restriction of the sample to young, remitted patients.

Cutting & Murphy (1988) compared subacute/chronic schizophrenics and control groups on a test of 'real world knowledge' – a series of simple multiple choice questions requiring the exercise of practical and social common sense. Three-quarters of the schizophrenic patients were found to score more than two standard deviations below control group means. The authors considered that this abnormality – which is at the very least closely related to semantic memory – was central to the various forms of thinking disturbance in schizophrenia, but its relationship to delusions was not specifically explored.

Tamlyn *et al* (1992) applied the best-known semantic memory task to a sample of schizophrenic patients encompassing all grades of severity and chronicity. In the 'silly sentences' test (Collins & Quillian, 1969) the subject is required to classify as true or false 50 statements like 'rats have teeth' and 'desks wear clothes'. Speed of verification is taken as a measure of the integrity of semantic memory (the verifications are all extremely easy, and errors are rarely made). For the schizophrenic patients, the mean verification time was highly significantly slower than for normal controls: even the fastest patients were slower than the average normal individual, and around two-thirds of the sample were outside the normal range altogether.

A further finding of this study was that nearly a quarter of the sample made multiple semantic errors, typically classifying obviously false statements as true. In contrast to the normal young or elderly who typically make 0–2 errors (Baddeley *et al*, 1987), 14 patients made three or more such errors, with five of these making more than 10. While semantic slowing and errors were significantly correlated with presence of formal thought disorder and with negative symptoms, no significant correlation with delusions was found. However, the correlations with individual symptoms were not strong and some patients who made substantial numbers of semantic errors had clinical pictures characterised principally by fantastic delusions.

In summary, two of the three studies which have examined semantic memory in schizophrenia have found preliminary evidence of a marked disorder. Although there are indications that this is relevant to thinking disturbance, whether it can be related to the clinical phenomenon of delusions remains to be established.

Conclusion

The two lines of argument developed here are certainly speculative and might be considered unconvincing. Before doing so, however, some points need to be borne in mind.

Firstly, Gray's (1982) theory of septo-hippocampal function, on which the animal behavioural line of argument draws heavily, should not be dismissed as just one theory among many. Rather, it is based on an even-handed review of the accumulated animal experimental findings and represents a closely argued integration of many earlier theoretical formulations. It might not be going too far to suggest that, while the septo-hippocampal system may not operate exactly as Gray argues, it is probable that it does not perform any vastly different function. Secondly, while it can be legitimately maintained that the cognitive psychological line of argument fails to relate semantic memory specifically to the phenomenon of delusions, it is difficult to disregard a growing feeling that there is a disorder of knowledge in schizophrenia. From Hemsley's (1987) and Gray *et al*'s (1991)

https://doi.org/10.1192/S0007125000296451 Published online by Cambridge University Press

weakening of the influences of stored regularities, through Cutting & Murphy's (1988) deficient real world knowledge, to Frith & Frith's (1992, in press) defect of second order representations, a number of theoretical accounts have invoked something very similar to abnormal knowledge as fundamental to the psychopathology of schizophrenia. Finally, both of the present lines of argument show striking points of similarity to Maher's (1974) theory of delusions, with the critical difference that, rather than being localised to its inputs, the abnormality is hypothesised to be in the cognitive machinery of belief itself.

References

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

BADDELEY, A. D. (1990) *Human Memory: Theory and Practice*. Hove: Erlbaum.

———, HARRIS, J., SUNDERLAND, A., *et al* (1987) Closed head injury and memory. In *Neurobehavioral Recovery from Head Injury* (eds H. S. Levin, J. Grafman & H. M. Eisenberg). New York: Oxford University Press.

——— & WILSON, B. (1988) Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition*, **7**, 212–230.

BERRIOS, G. E. (1991) Delusions as 'wrong beliefs': a conceptual history. *British Journal of Psychiatry*, **159** (suppl. 14), 7–14.

BLEULER, E. (1911) *Dementia Praecox, or the Group of Schizophrenias* (transl. 1950 by J. Zinkin). New York: International Universities Press.

COLLINS, A. M. & QUILLIAN, M. R. (1969) Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour*, **8**, 240–247.

CUTTING, J. & MURPHY, D. (1988) Schizophrenic thought disorder: a psychological and organic interpretation. *British Journal of Psychiatry*, **152**, 310–319.

FRITH, C. D. & FRITH, U. (1992) Elective affinities in schizophrenia and childhood autism. In *Social Psychiatry: Theory, Methodology and Practice* (ed. P. Bebbington). New Brunswick, New Jersey: Transactions.

GRAY, J. A. (1982) *The Neuropsychology of Anxiety: An Enquiry into the Function of the Septo-Hippocampal System*. Oxford: Oxford University Press.

———, RAWLINS, J. N. P., HEMSLEY, D. R., *et al* (1991) The neuropsychology of schizophrenia. *Behavioral and Brain Sciences*, **14**. 1–84.

HEMSLEY, D. R. (1987) An experimental psychological model for schizophrenia. In *Search for the Causes of Schizophrenia* (eds H. Hafner, W. F. Gattaz & W. Janzarik). Heidelberg: Springer Verlag.

JASPERS, K. (1959) Translated (1963) as *General Psychopathology* by J. Hoenig & M. W. Hamilton. Manchester: Manchester University Press.

KOH, S. D. (1978) Remembering of verbal materials by schizophrenic young adults. In *Language and Cognition in Schizophrenia* (ed. S. Schwartz). Hillsdale, New Jersey: Erlbaum.

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

MCKENNA, P.J. (1987) Pathology, phenomenology and the dopamine hypothesis of schizophrenia. *British Journal of Psychiatry*, **151**, 288–301.

SCHNEIDER, K. (1949) Translated (1974) as The concept of delusion. In *Themes and Variations in European Psychiatry* (eds S. R. Hirsch & M. Shepherd). Bristol: John Wright.

SQUIRE, L. R. (1987) *Memory and Brain*. Oxford: Oxford University Press.

SWANSON, L. W. (1978) The anatomical organization of septo-hippocampal projections. In *Functions of the Septo-Hippocampal System, CIBA Foundation Symposium 58* (eds K. Elliott & J. Whelan). Amsterdam: Elsevier.

TAMLYN, D., MCKENNA, P. J., MORTIMER, A. M., *et al* (1992) Memory impairment in schizophrenia: its extent, affiliations and neuropsychological character. *Psychological Medicine* (in press).

TULVING, E. (1983) *Elements of Episodic Memory*. Oxford: Oxford University Press.

WALKER, C. (1991) Delusion: what did Jaspers really say? *British Journal of Psychiatry*, **159** (suppl. 14), 7–14.

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

P. J. McKenna, MA, MB, ChB, MRCPsych, *Consultant Psychiatrist, Fulbourn Hospital, Fulbourn, Cambridge CB1 5EF*

https://doi.org/10.1192/S0007125000296451 Published online by Cambridge University Press