Varieties of confabulation and delusion

Michael D. Kopelman

King’s College London, Institute of Psychiatry, London, UK

Introduction. Different types of confabulation or false memory can arise from brain disease. There are competing explanatory theories for the mechanisms underlying confabulation. Recent literature has attempted to relate the notion of delusion to that of confabulation.

Method. A brief review of the literature relating to these ideas.

Results. The varieties of confabulation or false memory that can arise from brain disease are considered. The varieties of delusion and the contexts in which they arise are considered. Comparisons are made between the characteristics of spontaneous confabulation and those of delusional memory.

Conclusion. It is suggested that global theories purporting to account for both confabulation and delusions, in whatever circumstances they arise, can have only limited explanatory power. On the other hand, there are resemblances between confabulation and delusional memory, and the similarities and differences between these phenomena deserve further empirical investigation.

Keywords: Confabulation; False memory; Delusion; Delusional memory.

INTRODUCTION

In this paper, I will give a brief outline of the varieties of confabulation and delusion, which can arise in neurological or psychiatric disorders. Questions arise concerning whether false memories occurring in these different contexts represent the same or differing phenomena, and whether they reflect common underlying dysfunctions or deficits. In considering these issues, I will focus upon a comparison of “spontaneous confabulation” and “delusional memory”.

Correspondence should be addressed to Professor Michael Kopelman, Academic Unit of Neuropsychiatry, 3rd Floor, Adamson Centre, South Wing, St. Thomas's Hospital, Lambeth Palace Road, London SE1 7EH, UK. E-mail: michael.kopelman@kcl.ac.uk

© 2009 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business
http://www.psypress.com/cogneuropsychiatry
DOI: 10.1080/13546800902732830
CONFABULATION

Types of confabulation

Confabulation broadly refers to false or erroneous memories arising in the context of neurological disease. The memories may be either false in themselves or “real” memories jumbled in temporal context and retrieved inappropriately.

Following Berlyne (1972), I distinguished between “spontaneous” confabulation and “momentary” or “provoked” confabulation (Kopelman, 1987b). In spontaneous confabulation, there is a persistent, unprovoked outpouring of erroneous memories. By contrast, in momentary or provoked confabulation, fleeting intrusion errors or distortions arise in response to a challenge to memory, such as a memory test. Others have argued (e.g., Gilboa et al., 2006) that the distinction between spontaneous and provoked confabulations is not always clear-cut. For example, in normal conversation, a question might arise which provokes a confabulation—should this be called “spontaneous” or “provoked”? However, I would maintain that, difficult though it is to draw a precise boundary, the conceptual distinction is valuable—not only because the extreme forms of spontaneous confabulation are qualitatively quite distinct from simple intrusion errors or distortions, but also because there may have differing underlying mechanisms.

Further to this distinction between spontaneous and provoked confabulations, Schnider (Schnider, von Däniken, & Gutbrod, 1996; Schnider, 2003) attempted to tighten the definition of spontaneous confabulations by requiring that patients should have actually acted upon them, a tighter but more restrictive definition than Kopelman’s. Berlyne (1972) used the term “spontaneous” interchangeably with “fantastic”, whereas others have viewed bizarre or fantastic confabulations as a specific subtype of spontaneous confabulation, involving an additional dysfunction (Burgess & Shallice, 1996a).

Spontaneous confabulation

RJ’s confabulation was not limited to test situations. He confabulated in interacting with therapists, fellow patients, and his family. For example, one weekend while at home with his family he sat up in bed and turned to his wife, asking her, “Why do you keep telling people we are married?” His wife explained that they were married and had children, to which he replied that children did not necessarily imply marriage. She then took out the wedding photographs and showed them to him. At this point he admitted that the person marrying her looked like him but denied that it was he. (Baddeley & Wilson, 1986)
AB was a 43-year-old hospital employee, who was admitted to the Accident and Emergency department with Wernicke's encephalopathy (confusion ophthalmo-plegia, ataxia, and nystagmus), and she was found to have a half empty bottle of vodka in her handbag. Despite high doses of multivitamins, the patient continued to confabulate floridly, and subsequently cancer of the cervix was diagnosed. She confabulated in episodic memory, saying that she had been admitted for measles, and that her parents were visiting her regularly, despite the fact that they had been dead for 4 and 20 years respectively. She talked about being employed in the hospital, but identified the wrong hospital. She said that her brother was a doctor living on the 22nd floor, when her ward was on the top floor (the 12th). She also confabulated in semantic memory, saying that Stanley Baldwin was still the Prime Minister, and that Robert Maxwell (the newspaper proprietor) had been shot. (Kopelman, Ng, & van den Brouke, 1997)

BS was a 61-year-old man, who had been found collapsed on a pavement by a neighbour. He was confused and disorientated, and he showed pronounced nystagmus. He was also very ataxic, and he could stand only with support. At admission, he admitted to consuming a bottle of whisky a day. He was “stuck” in the 1970s or early 1980s, thinking that Margaret Thatcher was Prime Minister and Richard Nixon was President of the United States. He confabulated in episodic memory, thinking that he had recently taken a girlfriend home to his parents at Christmas, and that he was still serving in the airforce (as he had done in the 1970s). He also confabulated in personal semantic memory, thinking that his brother had emigrated to Canada, when the brother was in fact visiting on a weekly basis, and also in semantic memory, thinking that the Pope had recently been assassinated. (Kopelman, unpublished)

These three patients exhibited pronounced examples of spontaneous confabulation. Spontaneous confabulation most commonly occurs in autobiographical memory (Dalla Barba, 1993a; Dalla Barba, Cappelletti, Signorini, & Denes, 1997), but it can occur in semantic memory as well (Dalla Barba, 1993b; Kopelman et al., 1997; Moscovitch & Melo, 1997). Spontaneous confabulation has usually been attributed to executive dysfunction, resulting from frontal lobe pathology (Baddeley & Wilson, 1986; Kapur & Coughlan, 1980; Luria, 1976; Moscovitch & Melo, 1997; Stuss, Alexander, Liberman, & Levine, 1978), but confusional states and/or metabolic disturbance can also produce executive dysfunction and confabulation (DeLuca & Cicerone, 1991; Kopelman et al., 1997).

More recently, on the basis of lesion and/or functional imaging studies, spontaneous confabulation has been attributed more precisely to pathology within either ventromedial frontal cortex and orbitofrontal cortex (Gilboa et al., 2006; Gilboa & Moscovitch, 2002; Toosy et al., 2008; Turner, Cipolotti, Yousry, & Shallice, 2008) or in orbitofrontal cortex alone (Schnider, 2008; Schnider, Treyer, & Buck, 2005).
Theories of spontaneous confabulation can be broadly subdivided into those emphasising (1) context memory confusions or a source monitoring deficit; (2) trace specification/verification (strategic retrieval) deficits; (3) motivational theories, or (4) interactionist accounts.

**Context confusions/source monitoring deficits**

These accounts go back to Korsakoff (1889/1955, p. 404), who argued that many of his patients confused "old recollections with present impressions":

In telling of something about the past, the patient would suddenly confuse events and would introduce the events related to one period into the story about another period. ... Telling of a trip she had made to Finland before her illness and describing her voyage in fair detail, the patient mixed into the story her recollections of the Crimea, and so it turned out that in Finland people always eat lamb and the inhabitants are Tartars.

Korsakoff placed emphasis upon such temporal confusions in the genesis of confabulation, rather than on complete fabrications in memory, and subsequent writers in the clinical literature also made this observation (Moll, 1915; Talland, 1965; van der Horst, 1932; Victor, Adams, & Collins, 1971).

In a series of neat experiments, Schnider (Schnider, Ptak, von Daniken, & Remonda, 2000; Schnider et al., 1996) demonstrated that spontaneous confabulators (defined in terms of their having acted upon their confabulations) could be differentiated from other amnesic patients and from healthy controls on the basis of their errors on a temporal context memory task, but not on other memory and executive tasks. Moreover, this very specific deficit subsided as the confabulation improved. Although the experimental data on which he bases his argument remain essentially the same, Schnider (2003, 2008) now interprets this dysfunction more generally as a failure in "reality monitoring", resulting from the malfunctioning of a very rapidly acting (200–300 ms) filter, located within the orbitofrontal cortex, which sets the "cortical format" for subsequent memory encoding. However, Gilboa et al. (2006) found that other difficult discriminations (unrelated to temporal context) can also produce a high rate of false positive identifications by confabulators; and in preliminary observations, Bajo, Fleminger, and Kopelman (2008) showed that most (83%), but not all, confabulators produce the predicted pattern of performance on Schnider’s task.

Somewhat related is Dalla Barba’s (1993a; Dalla Barba et al., 1997) variant of this hypothesis. Dalla Barba has argued that so-called “temporal consciousness” is intact in confabulating patients, but it is malfunctioning. These patients are aware of a past, present, and future (unlike severely
amnesic patients), but, in making memory judgements, they employ only the most stable elements from their long-term memory stores. They tend to be “stuck” in the past, and this usually (but not always) means their early memories. Consequently, asked what they did today or what they will do tomorrow, confabulating patients reply with the well-established memories from their pasts, however irrelevant these memories may be to their present situation. Dalla Barba has investigated this deficit in patients with Korsakoff’s syndrome, other amnesic patients, and Alzheimer patients. (See also Dalla Barba & Boissé, 2010 this issue.)

Trace specification and verification deficits

Moscovitch (1989) postulated a deficit in the strategic component of retrieval processes, such that loose rules of plausibility and association are left to guide memory retrieval, and the resulting “memories” are not edited or suppressed in a normal fashion. Somewhat similarly, Conway and Tacchi (1996) postulated a failure in the executive evaluation of accessed memories during confabulation, and Kopelman (1987b, p. 1482) attributed spontaneous confabulation to “the extremely incoherent and context-free retrieval of memories and associations”. Burgess and Shallice (1996a) postulated deficits in a description process, an editor process, and a mediator process, which all contributed differently to the clinical phenomena of confabulations. The “descriptor” produced a specification of the type of trace that would satisfy the demands of a retrieval task, and “noisy” specification would increase the chance of an inappropriate representation being produced as a candidate memory. The “editor” checked that the output of a long-term storage system fitted with previously retrieved memory elements, and also with the overall task requirement. When the “editor” was impaired, confabulators would respond to a question without giving it adequate consideration or without checking and self-correcting themselves, somewhat similar to a previous hypothesis which had been proposed by Mercer, Wapner, Gardner, and Benson (1977). The “mediator” controlled cognitive (strategic and problem-solving) operations concerning the adequacy or plausibility of retrieved memory elements under strategic control. Impairment of this module resulted in reasoning errors and produced bizarre or fantastic responses.

Schacter, Norman, and Koutstaal (1998) proposed a general “constructive memory framework”, emphasising both encoding operations (initial binding of the distributed features of an episode together as a coherent trace with sufficient pattern separation of similar episodes) and retrieval processes (formation of a sufficiently focused retrieval description with which to probe memory stores, plus postretrieval monitoring and verification). If episodes
were not stored in a manner that allowed them to be accessed separately at retrieval, or if retrieval cues were not specific or focused enough, or if criterion setting were lax, memory distortion or confabulation would be likely to result.

Moscovitch and Melo (1997) postulated that confabulation was the product of impaired cue-retrieval (cues can be ambiguous, resulting in retrieval errors in normal subjects, but especially if memory is impaired), faulty strategic search (producing misleading cues and, thereby, inappropriate memories), and defective monitoring (the resulting errors would not be edited out). More recently, Gilboa et al. (2006) have modified this theory a little, suggesting that a failure in strategic retrieval and postretrieval monitoring, related to ventromedial and orbitofrontal pathology, is critical for spontaneous confabulation to arise. They argued that postretrieval monitoring has at least two components: an early, rapid, preconscious component, impairment of which is sufficient to cause confabulation, and a conscious elaborate monitoring of retrieval content for inconsistencies, conflicting evidence, and compatibility with task requirements. They argued that a failure to make fine-grained distinctions within memory could account for Schneider’s observations.

Finally Metcalf, Langdon, and Coltheart (2007), taking as their starting-point Langdon and Coltheart’s (2000) model of delusional belief, argued that two deficits are likely in most cases of spontaneous confabulation—an executive retrieval deficit and an evaluation deficit. In addition, they postulated that a specific personal bias may influence the content of confabulation.

**Motivational hypothesis**

Conway and Tacchi (1996) had earlier argued that a person’s current preoccupations and motivations can strongly influence the content of confabulation. Conway and Tacchi (1996, p. 333) gave a graphic account of how their patient used her confabulations to transform “the present into a time of harmony and comfort rather than distress” in a manner not dissimilar from Blanche Dubois in *A Streetcar Named Desire*. They explained the patient’s confabulation in terms of a combination of a failure of the executive editing of memories and of motivational biases.

Following this line of argument, Fotopoulou and colleagues (Fotopoulou, Conway, Griffiths, Birchall, & Tyrer, 2007; Fotopoulou, Conway, & Solms, 2007; Fotopoulou, Conway, Solms, Tyrer, & Kopelman, 2008; Fotopoulou, Solms, & Turnbull, 2004) have demonstrated that the content of confabulations are more likely to include pleasant experiences or positive self-representations than the corresponding “reality”. Confabulating patients are also more likely to identify incorrectly pleasant autobiographical memories as currently relevant. (See Fotopoulou, 2010 this issue; also Turnbull, Jenkins, & Rowley, 2004).
In this connection, Jaspers’s (1913/1974) distinction between the “content” or “theme” of abnormal mental phenomena (including motivational factors) and their “form” (relating to an underlying deficit or dysfunction) needs to be drawn.

Interaction accounts

Johnson and colleagues (Johnson, 1991; Johnson, Hashtroudi, & Lindsay, 1993) argued that confabulation might result from a wide range of source (context) or reality monitoring deficits (differentiating “real” memories from the imagined) which, in turn, could be secondary to impairments in encoding, retrieval, motivation, or judgement processes. However, Johnson, O’Connor, and Cantor (1997) failed to show that a severely confabulating patient differed from three other patients with frontal lesions in terms of performance on several measures of source monitoring. Hence, they concluded that confabulation might reflect an interaction between three factors—a vivid imagination, an inability to retrieve autobiographical memories systematically, and source monitoring deficits.

Somewhat similarly, Kopelman et al. (1997), following a detailed analysis of a severely confabulating patient’s errors, suggested that many confabulations (particularly in episodic memory) result from the conflation and inappropriate retrieval of “real” memory fragments out of temporal sequence, but that other confabulations result from perseverations (particularly in semantic memory), or from the patient giving instantaneous, ill-considered, and unchecked responses to immediate environmental and social cues.

Momentary confabulation

The fleeting intrusion errors or distortions made by patients in memory tests (e.g., story recall) resemble those made by healthy subjects in many contexts. As Bartlett (1932, 175 [1995 reissue]) wrote of story recall:

Epithets are changed into their opposites; incidents and events are transposed; names and numbers rarely survive intact . . . opinions and conclusions are reversed.

. . . At the same time, the subjects may be very well satisfied with their efforts believing themselves to have passed on all important features with little or no change. . . . Condensation, elaboration and invention are common features of ordinary remembering.

Consistent with this, Hammersley and Read (1986) found that healthy subjects were more vulnerable to interference effects from misleading information when recalling a story at 1 week’s delay than at immediate recall, and that
misleading information was most likely to produce errors if it was presented shortly before the recall of a story. Similar intrusions and distortions by both adults and children have been demonstrated in other experimental tasks (Lindsay & Read, 1994; Read & Lindsay, 1997; Schacter et al., 1998).

Kopelman (1987b) deliberately compared the recall of amnesic patients with that of healthy controls at a prolonged delay. The recall of Logical Memory passages by Korsakoff and Alzheimer patients at immediate recall and 45 minutes’ delay was compared with that of healthy controls at a 1-week delay. This design reduced the overall level of the healthy participants’ recall scores to nearer that of the amnesic patients, and 47% of the healthy participants showed some evidence of momentary confabulation at the 1-week delay compared with 50% of the Korsakoff patients and 44% of the Alzheimer group at immediate or 45 minutes’ delay. Moreover, there were qualitative similarities between the intrusion errors produced by each group. Consequently, Kopelman concluded that momentary or provoked confabulation in patients reflects “normal” mechanisms that occur when memory is “weak”.

Although “momentary confabulations” are usually a “normal” phenomenon arising when memory is poor, and do not imply any underlying pathology, Turner et al. (2008) have recently argued that, in patients with ventromedial or orbitofrontal pathology, both spontaneous and provoked confabulations can arise.

Summary

In summary, spontaneous confabulation can be profuse, bizarre, preoccupying, and held with absolute conviction. Spontaneous confabulations appear to result from ventromedial or orbitomedial neuropathology. They may reflect faulty trace specification and verification (monitoring) of retrieved memories, confusions in temporal context, perseverations especially in semantic memory, an unchecked response to social or environmental cues, and/or underlying motivational and emotional biases. By contrast, momentary or provoked confabulations are fleeting distortions or intrusion errors. These can be seen in healthy subjects when their memory is “weak”, and so do not necessarily imply any underlying pathology, although Turner et al. (2008) have recently suggested that ventro- or orbitomedial frontal pathology can also give rise to these confabulations.

OTHER FORMS OF FALSE MEMORY

Kopelman (1999) described other instances in which false memory can arise, some of which are not necessarily related to overt neurological damage.
The false recognition syndrome

Patients with frontal lesions have been described who show a high rate of false recognition responses on cognitive testing, not necessarily related to “extensive” spontaneous confabulation (Delbecq-Derouesné, Beauvois, & Shallice, 1990; Parkin, Bindschaedler, Harsent, & Metzler, 1996; Schacter, Curran, Galluccio, Milberg, & Bates, 1996). In two of these patients (Parkin et al., 1996; Schacter et al., 1996), the pathological false recognition appeared to be particularly evident when novel material was presented from previously studied semantic categories, but it was also present in response to nonstudied words, pseudowords, faces, and environmental sounds. It was suggested that false alarms arose when the test items were “generally” (Schacter et al.) or “incidentally” (Parkin et al.) consistent with the class, category, or characteristics of a study list, and that the patients failed to recollect the identity of particular items. In such instances, both patients seemed to show an overreliance on familiarity judgements. However, unlike confabulating patients, these patients showed a relative preservation of recall for everyday experiences. It remains unclear whether such cases provide a “model” for one particular component of spontaneous confabulation, or whether theirs is a qualitatively distinct disorder.

Recovered memory for childhood sexual abuse

This hugely controversial topic has been extensively reviewed elsewhere (e.g., Conway, 1997; Read & Lindsay, 1997). It has been claimed that many apparently recovered memories of childhood sexual abuse are likely to be false, particularly if retrieved after a long delay or in response to leading questions by therapists. Although cases of forgetting of traumatic experiences, including childhood sexual abuse, have certainly been reported (Loftus, Polonsky, & Fullilove, 1994), and empirical evidence of recovery of such memories has also been described (Brewin, 2007; Schooler, Ambadar, & Bendiksen, 1997), it seems likely that at least some recovered memories are false. Shimamura (1997) documented evidence that highly emotional or traumatic experiences can enhance memory storage transiently but in a fragmented form, in which “free floating” memory fragments are poorly located in temporal and spatial context, particularly in the absence of memory rehearsal. On the basis of what is known about frontal and medial-temporal mechanisms in memory, Shimamura argued that such memory fragments would be especially vulnerable both to forgetting and to distortions and augmentations such that, if memory were reinstated, it would be very unlikely to be either accurate or complete. A somewhat similar argument was proposed by Schacter (1996; Schacter, Norman, & Koutstaal,
1997). In addition, the genesis of such erroneous memories almost certainly cannot be understood independently of the social contexts in which they arise (Kopelman, 1997).

Confabulation in schizophrenia

Nathaniel-James and Frith (1996) described “confabulation” in schizophrenic patients. They required their patients to recall a fable/story, and then to identify the correct moral from it. They found a significantly increased rate of intrusion errors or confabulations in schizophrenic patients’ recall compared with healthy participants. This remained true even after attempting to control for the level of recall performance, and it appeared to be correlated with performance on an executive test (the Hayling; Burgess & Shallice, 1996b). The authors argued that these patients’ confabulations were qualitatively different from those made by healthy subjects in that they tended to introduce irrelevant or unrelated material into the stories, and, in a subsequent investigation, Nathaniel-James, Foong, and Frith (1996) showed that the schizophrenic patients were unable to distinguish plausible from implausible stories or to correct their errors after listening to a tape-recording of their own recall. However, Kopelman (1999) argued that it remained possible that the “confabulations” of schizophrenic patients are simply momentary intrusion errors and distortions, resulting from the impaired memory which has been previously described in such patients, coloured by their preoccupying delusional beliefs (see later).

False confessions

Gudjonsson (2003) has described the various circumstances in which people come to make false confessions on a voluntary basis (see also Gudjonsson & MacKeith, 1988), which he labels “confabulations”. For example, Gudjonsson, Kopelman, and MacKeith (1999) described a case of false confession in a 17-year-old man who was eventually released from prison 25 years later. He was a man of low self-esteem and high suggestibility who, during the course of a police interview lasting 48 hours (in the absence of a lawyer or doctor) in which he was in a distressed and aroused state, progressed from thinking “It might have been me” to “I don’t know if I killed her or not. I keep seeing her”, through “I must have done it because I can see a picture of her” to “I am sure I killed her . . . I know I did it.” One of the factors that contributes to false confessions is source amnesia (Johnson et al., 1993), which gives rise to a form of “memory distrust”, such that interviewees become confused concerning whether “memories” have been generated internally or from external sources of information (source memory errors).
The false confessions which result may then become internalised, in which case Gudjonsson describes them as “confabulations”; in a series of investigations, Gudjonsson and colleagues have investigated the circumstances and correlates of such false confessions.

**Pseudologia fantastica**

Occasional patients are seen who create a web of fantasies, lies, and untruths around themselves, almost compulsively, and Fish (1967) classified this phenomenon as a form of confabulation. This tends to occur in people who are of low IQ or self-esteem, but who do not have known brain disease or frontal lobe dysfunction. Kopelman (1999) reported a young man, poorly educated and not very bright, who talked endlessly about being a pop star, about the other stars he knew, and about his show business activities, despite the fact that it soon became clearly apparent that he had led a fairly mundane life in South London. He had given himself a derivation of a famous pop star’s name, and he claimed that he was “world-wide famous in Streatham”. Such activity is not necessarily harmless. In the case of R. v. O’Brien and others, 1999 (also known as the “Cardiff 3”), the Court of Appeal established that Mr H’s confession to the murder of a milkman 12 years’ earlier (which had resulted in his imprisonment and that of two acquaintances) was likely to have been a fabrication. Mr H had lied about many things during his childhood and adolescence, claiming (for example) to have been a successful rugby football player (he was crippled following a series of operations for two club feet). One of the most persuasive pieces of evidence in court came from his former solicitor, who described how Mr H had confessed to a previous offence, which he could not have done because he was being interviewed by the police for something else at the time that that alleged offence had taken place!

**Summary**

In brief, false memories can arise in many situations, which are not necessarily related to overt neurological damage. Hence, Kopelman (1999) put forward a model which incorporated input from the social environment. It also incorporated a personal semantic belief system which biased retrieval from autobiographical memory, into which I have now incorporated prevailing preoccupations, mood state, and personal fantasy to take account of the motivational biases within a person’s confabulation (Figure 1). Some forms of false memory are a direct consequence of neurological damage, e.g., spontaneous confabulation as a result of damaged control or “filter” systems within the ventromedial and orbitomedial frontal regions, or the false
recognition syndrome arising from dysfunction in Schacter et al.’s (1996) postulated prefrontal/cerebellar error detection circuits. Other forms of false memory derive from a combination of psychosocial input, prevailing emotional preoccupations, mood state, and/or personal fantasies biasing the frontal control mechanisms, and giving rise to, e.g., source memory errors and internalised confabulations in false confession or the “confabulations” of schizophrenia.

**DELUSIONS**

A delusion is a false belief, held as an absolute conviction, not amenable to argument, not culturally explicable, often bizarre, and usually preoccupying (Clare, 1976; Kopelman, 1994; Mullen, 1986).

The content and character of delusions may take many forms (Cutting, 1997; Gelder, Gath, & Mayou, 1983; Mellor, 1970; Mullen, 1986; Schneider, 1959). Among the so-called first-rank symptoms of schizophrenia are delusions of control (relating to actions, affect, or volition) and delusions concerning thoughts (thought insertion, thought withdrawal, or thought diffusion/“broadcasting”). Grandiose delusions may occur in mania or in frontal lobe disease. By contrast, depression may give rise to delusions of worthlessness or guilt, hypochondriacal/somatic delusions, or nihilistic
delusions including Cotard’s syndrome. Religious delusions may occur in a
variety of syndromes, including epilepsy. Delusions of reference and persec-
utory delusions may occur in many psychiatric and neurological disorders,
including alcohol- or substance- induced psychoses. Sexual, amorous, or
jealous delusions also occur in various contexts, particularly alcohol abuse.
Delusions of being burgled are characteristic of dementia. Delusions of
misidentification, e.g., the Capgras syndrome, can occur in either neurological
or psychiatric disorder; in the former, they tend to be associated with
combined right parietal and frontal pathology (Alexander, Stuss, & Benson,
1979; Ellis & Young, 1990).

Furthermore, delusions arise in differing circumstances (Cutting, 1997;
Mullen, 1986; Schneider, 1959). For example, delusions can occur
apparently “out of the blue” (primary or autochthonous delusions), or
they may emerge from a particularly perplexed psychological state (known
as “delusional mood” or “delusional atmosphere”), in which a sense of
perplexity, foreboding, dread, and anxiety precedes the occurrence of frank
delusions. Alternatively, delusions can arise secondarily to a “real”
perception (as in the misleadingly called “delusional perception” or in
somatic passivity experiences), or as a consequence of hallucinations (e.g.,
auditory hallucinations expressing derogatory or persecutory ideas).
Delusions can also emerge secondarily to an affective state (mania,
depression), or in consequence of an acute medical disorder (e.g.,
confusion, delirium tremens) or chronic neurological disease (e.g.,
temporal lobe epilepsy, dementia). As already mentioned, there is a somewhat
different flavour to delusions in these different situations, and any
comprehensive theory of delusions needs to account for this.

Cognitive theories of delusions will be the major topic of some of the
other papers in this volume. Amongst these theories, Maher (1992) has
argued that delusions or false beliefs arise as rational responses to
abnormal perceptual experiences. Ellis and Young (1990; Ellis, Young,
Quayle, & de Pauw, 1997) attributed the Capgras delusion to disrupted
signalling of the emotional or affective significance of a face in the two-
route model of face processing (Bauer, 1984; Bruce & Young, 1986); and
they attributed the Fregoli delusion to dysfunction in the person identity
nodes and the cognitive system of that model. Similarly, Young,
Robertson, Hellawell, de Pauw, and Pentland (1992) attributed the Cotard
delusion to an underlying lack of feelings of familiarity, resulting in a
sense of unreality (compare Leafhead & Kopelman, 1997). Frith and
Corcoran (1996) attributed persecutory delusions in part to abnormal
“theory of mind” experiences, and Bentall (2003) to an abnormal pattern
of attributions, protecting the individual against chronic feelings of low-
self esteem. The latter might also relate to Garety and Freeman’s
(1999) finding that psychotic patients tend to jump to conclusions on
probabilistic reasoning tasks. Coltheart and colleagues (e.g., Davies & Coltheart, 2000; Langdon & Coltheart, 2000; McKay, Langdon, & Coltheart, 2005, 2007) have argued in favour of (1) a primary perceptual/affective deficit, which distorts or makes anomalous current sensory information, semantic or autonomic input, or emotional experience; and (2) an abnormality in belief evaluation, which causes the uncritical acceptance/maintenance of an implausible hypothesis or belief. In addition, there may be personal attributional biases which nuance the favoured explanations that are generated to account for the anomalous experience.

There are two points that can be made here. First, it is not clear that any of these theories can account for the entire variety of delusional beliefs, the differing circumstances in which they can arise, and for why particular delusions characterise specific disorders/age groups. In order to explain such phenomena, the present very general theories have to be stretched in diverse directions. Bell, Halligan, and Ellis (2006) have made a related point, whilst distinguishing between theories which explain delusions as a breakdown of normal belief formation, those which explain only the pathology, and those approaches which view delusions as one end of a continuum of anomalous mental phenomena. Figure 2 attempts to incorporate some of these factors into the Langdon and Coltheart (2000) model. Hence, it indicates that the prevailing psychological state, hallucinations, mood state (compare Cutting, 1997; Mullen, 1986; Turnbull & Solms, 2007), or somatic disease can precipitate delusion formation, whether by promoting a perceptual anomaly and/or by producing a deficit in belief evaluation. Likewise, the age and sex of an individual, not to mention his/her underlying clinical condition (as discussed above) can influence the content and character of the specific delusions which arise. This model leaves to be clarified whether delusions always result from a perceptual anomaly or whether they can occur “out of the blue” or secondary to a “real/true” perception, as traditional theory postulates (Jaspers, 1923/1963; Schneider, 1959).

Second, unlike confabulation, delusions are not in themselves necessarily a memory phenomenon, although they become incorporated into memory. For example, relatively common delusions—for instance, that one’s thoughts are being controlled or that one has a special religious mission, or that spies are lining up outside the house—do not necessarily have a memory component, but instead may result from an anomalous perceptual experience, abnormal attentional biases, or from a dysfunctional belief evaluation system.

By contrast, delusional memories clearly are a memory phenomenon, although it should be noted that delusions are common in psychiatric patients whereas delusional memories are rare.
DELUSIONAL MEMORIES

Definitions

In the literature on psychiatric phenomenology, differing definitions of delusional memory are given, well summarised by Buchanan (1991). For example, Mullen (1986) described delusional memory as “a delusional insight [which] occurs not as an intuition about the world or as a change in...
knowledge of or about the world but in the form of a memory”. Buchanan argued that this definition of delusional memory is analogous to Jaspers’s (1923/1963) concept of a “delusional idea”. By contrast, Gelder et al. (1983) stated that: “It is not the memory that is delusional but the interpretation that has been applied to it.” Buchanan (1991, 472) stated that this definition is much more analogous to Jaspers’s (1923) concept of “delusional perception” (a “real” perception giving rise to delusional interpretation). Buchanan pointed out that Schneider (1959) incorporated both definitions: “If a special meaning is subsequently attached to a remembered percept, then this becomes a delusional perception with two component parts. . . . A similar belief can also be held as a delusional intuition [idea] if, for example, it comes into someone’s head that he [she] had supernatural gifts as a child.”

In other words, a delusional memory can be akin to either a delusional perception or a delusional idea (intuition).

This duality of description led Kopelman (1997) to define delusional memory as consisting of either a true memory that gives rise to a deluded interpretation or (perhaps more commonly) a false memory arising in the context of a psychosis.

Case examples

DL was a 33-year-old man, first seen in 1981, 7 years after a high profile murder in London. Lord Lucan had disappeared after his children’s nanny was found battered to death at the family’s Belgravia home, and Lady Lucan had also been viciously beaten up. To this day, Lord Lucan has never been found, although sightings have been reported intermittently in various parts of the world. DL gave a graphic account of having been hired to carry out the attempted killing of Lady Lucan, claiming that he had been introduced to Lord Lucan at a nightclub, where the plan had been put to him and Lady Lucan had been pointed out to him. He recalled that he had indeed assaulted Lady Lucan, and that he had killed both Lord Lucan and the couple’s nanny, and he described in detail both the meeting at which he had been hired and the killings. The account was so convincing that I was wondering about telephoning the police until the patient told me about the “angels” on the bonnet of his car. There was no evidence of any neurological or medical condition, but this “memory” occurred in the context of other “first-rank” symptoms of schizophrenia (the patient’s first episode). Following treatment with a neuroleptic medication (thioridazine), these delusional preoccupations were substantially ameliorated, although, as commonly occurs in such cases, the “memory” was never completely abolished. (Kopelman, 1997)

WM was a 47-year-old clerical worker, who was studying part time for a PhD in English Literature, and she was extremely articulate about Chaucerian and Shakespearian literature. However she claimed that, in 1970, she had been working on a fruit-picking farm during the summer in East Anglia, when she encountered an
internationally famous orchestral conductor, who also happened to be fruit-picking there. No words were exchanged between them, but she claimed that, subsequently, he traced her, and he followed her in London, and that her friends had challenged him about this. She reported that she retreated to a rural town, where her parents were then living, but the musician followed her there. She believed that he was in love with her, and she said that she was prepared “to meet him half way”. She claimed that he had recently been divorced, and he was psychologically unsettled at that time. Subsequently, he stopped pursuing her, and they have exchanged only a few words since, when she waited for him outside stage doors. However, she had written to him on a regular basis, although never receiving a reply. On one occasion, he was sent a final demand by an expensive London store after she had purchased a wedding dress and had arranged for the bill to be sent to him. On another occasion, she deposited her suitcases outside his flat, intending to move in. Although being of apparently very sober, sensible demeanour, this patient also described “first-rank” symptoms of schizophrenia, including thought insertion, thought diffusion, and controlled actions. All her abnormal beliefs appeared to derive from her “memory” of meeting the conductor on a fruit farm, and there is no evidence that that meeting ever took place. (Kopelman, Guinan, & Lewis, 1995)

FJ, an Irishman, had been 39 when he was charged with killing his employer and friend in an arson attack at the friend’s newspaper store in 1976. The testimony against him came from two Irish acquaintances, who were arrested and charged with the murder shortly after it happened, and who turned evidence against him. When seen by psychiatrists 3 months before the trial, FJ was described as truculent, irritable, and difficult but not psychotic and fit to plead. However, during the course of the trial, FJ started to believe that the prison officers were poisoning his drinking water, that gas was being piped through a hole in the wall into his Old Bailey cell, and that there was a conspiracy against him involving the prosecution lawyers, the Judge, and ultimately his own lawyers. Consequently, FJ fired his own lawyers, carried out his own defence and, perhaps not surprisingly, he was convicted of the murder. Twenty-five years later, FJ had not developed any new psychotic symptoms in the intervening period, but he retained his abnormal beliefs as absolute convictions, and they had now become delusional memories. His conviction was overturned as “unsafe” in 2002 when he was 65, i.e., 26 years after his arrest. (Kopelman, unpublished)

Patient P was a 24-year-old PhD student at a well-known university. He had been abused by his stepfather when aged 12, and the stepfather was subsequently convicted of abusing someone else. During his teenage years and early twenties, P coped with the situation by consuming lots of alcohol and cannabis. P had had one brief homosexual relationship and a number of girlfriends. Following a difficult time in his life (involving the break-up of a relationship and unforeseen difficulties in his research which were not P’s fault), he started to feel very stressed, and he went to see a staff counsellor. He talked about the abuse in early adolescence, and about his alcohol and cannabis consumption. He then started to talk about his father having abused him, and he confronted his father about this. Over the next few days,
the accusations multiplied, involving claims of sexual abuse by his aunt, mother, and grandparents. A few days after this, P became overtly psychotic, believing that his therapist was in a conspiracy with his father. He was treated with a neuroleptic (trifluoperazine), and his delusional beliefs subsided over the next 2 to 3 weeks. At that point, there were still dim “memory fragments” but, eventually, he came to believe that only his stepfather had abused him, and indeed there was no evidence of abuse by any other relative. It appeared that these delusional memories emerged as an elaboration of one “real” memory in the context of a psychotic disorder, possibly alcohol and cannabis precipitated. (Kopelman, unpublished)

Delusional memory and spontaneous confabulation

In three of the above examples (DL, WM, P), the delusional memories appeared to arise spontaneously as false memories in the context of psychosis. In the fourth example (FJ), the delusional memories were not “true memories giving rise to a deluded interpretation” but consisted of false memories of “events” still being given a deluded interpretation (i.e., that they really did happen). Although it might be argued that, in this sense, any memory of an “old” delusion could be interpreted as a delusional memory, the difference is that FJ was still experiencing these memories with absolute (“delusional”) conviction and the desire to act upon them, whereas he was no longer experiencing any other hallucinations or delusions in his current everyday life, and he had not done so since his conviction 25 years earlier.

In any event, delusional memory is by definition a memory phenomenon, and it is the closest analogue in psychiatric patients of spontaneous confabulation in neurological cases. By definition, a delusional memory is held as an absolute conviction, not amenable to argument, and not culturally explicable. Moreover, delusional memories are usually intensely preoccupying and often bizarre. Similar features have been described in spontaneous confabulation by Baddeley and Wilson (1986), who pointed to closely similar characteristics in the confabulations of a patient with bilateral frontal damage following a traumatic head injury. Hence, the clinical phenomena of delusional memories can closely resemble those of spontaneous confabulations.

However, there appear to be certain respects in which delusional memories differ from spontaneous confabulations. The first is that they may be more elaborate and systematised, and sometimes they occur in people who are otherwise functioning normally without obviously abnormal behaviour (Baddeley, Thornton, Chua, & McKenna, 1996; Kopelman, 1999; Kopelman et al., 1995). This was true of the first two examples given earlier, and also of the third man when he was seen in 2002. Second, but related to this first point, delusional memory patients are often cognitively intact, whereas patients exhibiting spontaneous confabulation almost invariably...
show impairment on standard tests of executive function and are often profoundly amnesic. In the Kopelman et al. (1995) case, WM had a verbal IQ of 131, and she did not show any evidence of executive or anterograde memory impairment. Patient P had a NART-R IQ of 113, and he scored in the normal or superior range on Logical Memory, immediate and delayed, the Recognition Memory Test for words and faces, the Autobiographical Memory Interview, the Modified Card-Sorting test, cognitive estimates, and Trailmaking B. Similarly, David and Howard (1994) reported that four patients with delusional memories showed normal performance on the cognitive estimates test (Shallice & Evens, 1978); and Baddeley et al. (1996) concluded that the delusional memories and delusional beliefs of schizophrenic patients, although often bizarre, did not simply reflect a breakdown in executive function or autobiographical memory.

In summary, although delusional memories share many characteristics with spontaneous confabulation, they differ in that delusional memories appear to be thematic and unrelated to executive dysfunction, whereas confabulations are usually fluctuating and multifaceted as well as virtually always related to executive dysfunction (Baddeley et al., 1996; Kopelman et al., 1995; but see Burgess & McNeil, 1999).

However, the similarities and differences between delusional memories and spontaneous confabulation deserve further investigation. Moreover, although I am arguing that they should (at least for the present) be kept conceptually distinct, it must be acknowledged that, occasionally, certain specific forms of delusion, such as delusional misidentification, can coincide with or evolve from spontaneous confabulations, because of common underlying (frontal) pathology (Benson & Stuss, 1990; Box, Laing, & Kopelman, 1999; Mattioli, Miozzo, & Vignolo, 1999).

**CONCLUSION—THE REQUIREMENTS FOR THEORY**

Any comprehensive theory of false memories needs to take account of the very different situations in which false memories arise. Most existing theories focus purely upon spontaneous confabulation in brain disease. Kopelman (1999) considered various tentative possibilities. In spontaneous confabulation, ventromedial and orbitofrontal damage affect trace specification and verification systems, whereas momentary confabulation appears to result from a “normal” response to a “weak” memory trace. In delusional memory, Kopelman suggested that there may be anomalous processing of input modulated by personal self-beliefs, mental state, and expectations. In false confession, source memory errors occur, usually in the context of low self-esteem and/or depression and social coercion. In pseudologia fantastica, there may be anomalous, biased, or selective retrieval from autobiographical
memory. Some of these phenomena may arise from a combination of factors, e.g., the absence of rehearsal and a particular social context in cases of false memories for child abuse. In others, an interaction between social and biological factors may be important, e.g., the confabulations produced by brain-damaged patients in very stressful or extreme situations. These possibilities were interpreted within a very general model of memory and executive function, where social factors and a notion of “self” were incorporated (Kopelman, 1999), but they need to be explored further. Cognitive models need to accommodate the full range of contexts in which false memories can arise.

Second and similarly, any theory of delusions needs to take account of all the circumstances, primary and secondary, in which delusions arise, and the fact that they can have very different characteristics and content across these different contexts. It is not entirely clear that the existing models are able to do this, except in very general terms, but Figure 2 attempts to incorporate (in a preliminary fashion) some of the pertinent factors that can precipitate a delusion and/or determine its particular content or theme.

Third, confabulations, delusions, and delusional memories, in my view, need to be kept conceptually distinct. Although confabulations and delusional memories are, by definition, memory phenomena, other delusions may primarily reflect abnormal perceptual, attentional, or interpretative (belief) phenomena, even though they come to be incorporated into memory subsequently.

Although spontaneous confabulation and delusional memories share many characteristics, they differ in that confabulations tend to be fluctuating and multifaceted, as well as virtually always related to executive dysfunction, whereas delusional memories appear to be thematic and unrelated to executive dysfunction. However, specifying the characteristics and underlying psychopathology of these two memory phenomena remains an issue for further empirical investigation.

REFERENCES


