

---

# 1 What Is Confabulation?

They talk freely in the intervals of mitigation, but of things that do not exist; they describe the presence of their friends, as if they saw realities, and reason tolerably clearly upon false premises.

—John Coakley Lettsom (1787)

## 1.1 Introduction

A neurologist enters a hospital room and approaches an older man sitting up in bed. The neurologist greets him, examines his chart, and after a brief chat in which the man reports feeling fine, asks him what he did over the weekend. The man offers in response a long, coherent description of his going to a professional conference in New York City and planning a project with a large research team, all of which the doctor writes down. The only problem with this narration is that the man has been in the hospital the entire weekend, in fact for the past three months. What is curious is that the man is of sound mind, yet genuinely believes what he is saying. When the doctor informs him that he is mistaken, he replies, “I will have to check with my wife about that,” then seems to lose interest in the conversation. The man isn’t “crazy” or schizophrenic; he is quite coherent and can answer all sorts of questions about who his children are, who the current president is, and so on. He is *confabulatory*, owing in this case to the fact that he has Korsakoff’s syndrome, a disorder that affects his memory, producing a dense amnesia for recent events. But unlike other patients with memory dysfunction, who freely admit their memories are poor, a patient with Korsakoff’s syndrome will confidently report as memories events that either did not happen (or at least did not involve him) or that happened to him, but much earlier in life. This man’s act of describing the conference in New York City is known as a *confabulation*.

The neurologist moves down the hall to a room in which an older woman patient is in bed talking to her daughter. When the daughter sees him, she asks to speak with him outside in the hallway. “She won’t admit she’s paralyzed, Doctor. What’s *wrong* with her?” The woman has a condition familiar to anyone who has worked with people after a stroke, known as denial of paralysis (hemiplegia). The more general name for this condition is *anosognosia*, which means lack of knowledge about illness. It can come about when a stroke damages a certain part of the right hemisphere just behind and above the right ear, causing paralysis or great weakness on the left side of the body. The denial tends to occur right after the patient recovers consciousness, and tends to last only a few days.

The doctor walks back in the room, approaches the woman, and greets her. When he asks her how she is, she reports feeling fine. “Are

both your hands equally strong, Mrs. Esposito?" he asks. "Yes they're fine," she replies. "Can you touch my nose with your right hand?" he asks. She reaches up, a bit unsteadily, but succeeds in touching the doctor's nose. "Would you touch my nose with your left hand?" he then asks. Mrs. Esposito pauses a moment, rubs her left shoulder and replies, "Oh, I've got severe arthritis in my shoulder. You know that Doctor; it hurts." Again, she is not lying or pretending, she genuinely believes that she can move her arm. She also believes her confabulation about arthritis.

Perhaps what is most troubling about witnessing such confabulations is the rock-jawed certainty with which they are offered up. The patients give none of the outward appearances of lying, and indeed most writers on the subject do not consider confabulation to be lying, because it lacks at least two crucial components: the intent to deceive, and knowledge contrary to what is claimed. The claims about arthritis or the conference in New York City were not intentionally concocted with the motive of deception in mind; the patient is reporting what seems true to him or her.

Why then does confabulation happen? Confabulation seems to involve two sorts of errors. First, a false response is created. Second, having thought of or spoken the false response, the patient fails to check, examine it and recognize its falsity. A normal person, we want to say, would notice the falsity or absurdity of such claims. The patient should have either not created the false response or, having created it, should have censored or corrected it. We do this sort of censoring in our normal lives. If I ask you whether you have ever been to Siberia, for instance, an image might appear in your mind of you wearing a thick fur coat and hat and braving a snowy storm, but you know that this is fantasy, not reality. In very general terms, the confabulating patient lacks the ability to assess his or her situation, and to either answer correctly, or respond that he or she does not know. Apparently, admitting ignorance in response to a question, rather than being an indication of glibness and a low level of function, is a high-level cognitive ability, one that confabulators have lost. "I don't know," can be an intelligent answer to a question, or at least an answer indicative of good cognitive health.

Confabulation was initially considered solely a disorder of memory in which a patient gives false or contrived answers to questions about his or her past, but believes those answers to be true, as in the case of the man with Korsakoff's syndrome. Confabulation as a technical term was applied first to Korsakoff's patients by the German psychiatrist Karl Bonhoeffer in the early 1900s (Berrios 1998). In this narrower conception, confabulation occurs when patients produce stories that fill in gaps in their memories. The American Psychiatric Association's official *Diagnostic and Statistical Manual of Mental Disorders* (known as DSM IV), for instance, defines confabulation as "the recitation of imaginary events to fill in gaps in memory"

(1994, 433). However, confabulation also appears in a wide variety of other syndromes, many of which involve no obvious memory problems, including as we saw, anosognosia for hemiplegia (denial of paralysis), but also split-brain syndrome, Anton's syndrome (denial of blindness), Capgras' syndrome (the illusion that an impostor has replaced a person close to the patient), and schizophrenia.

The apparent diversity of confabulation syndromes invites a search for something they have in common. If a single brain region is damaged in all of these patients, and we know something about the function of that part of the brain, perhaps this knowledge can allow us to begin to unravel the mystery of why people confabulate and what it tells us about brain function. Unfortunately, it is not that simple because the sites of damage in confabulating patients seem to be widely scattered throughout the brain. Theories of the sites of lesions in confabulation have had two rough areas of focus: the frontal lobes and the corpus callosum—the large bundles of nerve fibers that interconnect the left and right hemispheres (Pandya and Seltzer 1986). Most recently, accounts of the locus of lesion in confabulation have tended to focus on the frontal lobes (Stuss et al. 1978; Benson et al. 1996; Burgess and Shallice 1996; Johnson et al. 1997). These memory-based frontal theories are examined in chapter 3. It has long been suspected, however, that the neural locus of confabulation may have an important lateral component, owing to the presence of confabulation in split-brain patients (see chapter 7). The fact that denial of paralysis happens overwhelmingly with right hemisphere strokes also seems to indicate a lateral component in confabulation (see chapter 6).

Confabulation has also been reported in young children reporting their memories, in subjects of hypnosis, and in normal people in certain experimental settings. When normal people are asked about certain choices they made, they can produce something that sounds rather like a confabulation. Nisbett and Wilson (1977) set up a table in a department store with pairs of nylon stockings and asked shoppers to select the pair they preferred. Unbeknown to the shoppers, all of the pairs were identical. People tended to choose the rightmost pair for reasons that are not clear, but when asked the reason for their choice, the shoppers commented on the color and texture of the nylons. When they were told that the nylons were identical, and about the position effects, the shoppers nevertheless tended to resist this explanation and stand by their initial reasons. As with patients with neurological disease, the question that is raised by such behavior is, why didn't the shoppers reply that they didn't know why they preferred that pair of nylons?

Rather than being merely an odd neurological phenomenon, the existence of confabulation may be telling us something important about the human mind and about human nature. The creative ability to construct

plausible-sounding responses and some ability to verify those responses seem to be separate in the human brain; confabulatory patients retain the first ability, but brain damage has compromised the second. One of the characters involved in an inner dialogue has fallen silent, and the other rambles on unchecked, it appears. Once one forms a concept of confabulation from seeing it in the clinic or reading about it in the neuropsychological literature, one starts to see mild versions of it in people. We are all familiar with people who seem to be unable to say the words “I don’t know,” and will quickly produce some sort of plausible-sounding response to whatever they are asked. A friend once described a mutual acquaintance as “a know-it-all who doesn’t know anything.” Such people have a sort of mildly confabulatory personality, one might say. One soon learns to verify any information they offer, especially if it involves something important.

One way to connect these normal cases of confabulatory people with the clinical examples is the idea that the normal people may be the same sort of people who exhibit clinical-level confabulation after brain injury. In a study of the premorbid personality of his anosognosic patients, Weinstein reported that “relatives saw them as stubborn, with an emphasis on being right” (1996, 345). Confabulation in the clinic might be produced by the suddenness of the injury, but on the normal spectrum, there may be all sorts of mild degrees of it among us. Those with clinical confabulation slowly learn to check, doubt, and finally inhibit their confabulations; similarly, normal people may become mildly confabulatory for a period as they age, then learn how to correct for it.

Confabulation involves absence of doubt about something one should doubt: one’s memory, one’s ability to move an arm, one’s ability to see, and so on. It is a sort of pathological certainty about ill-grounded thoughts and utterances. The phenomenon contains important clues about how humans assess their thoughts and attach either doubt or certainty to them. Our expressions of doubt or certainty to others affect how they hear what we say. In the normal social milieu, we like people to be certain in their communication, and strong social forces militate against doubting or pleading ignorance in many situations. A cautious weather forecaster, for example, who, when asked about tomorrow’s weather always replies, “I don’t know,” or “Can’t be sure,” will soon be unemployed. Excessive caution is especially counterproductive when it occurs in someone in a position of power or responsibility. Imagine a general who was never sure what to do, and as a consequence never did anything, never instructed his soldiers because he always felt that he didn’t know for certain what the right strategy was. Armies (and life forms) that do not move forward aggressively are soon overtaken by their enemies or competitors. Those under the authority of the leader can find admissions of ignorance especially troubling, even frightening. Imagine the president being asked what he plans to do

about a current oil crisis, for instance, and his answering, “I don’t know what to do about it.” The stock market would plummet. At least in some contexts, then, an answer that is possibly (or even probably) wrong is better than none at all.

There is also a clear connection here to the human gift for storytelling. Many confabulations are plausible little stories, about what one did over the weekend, or why one can’t move one’s arm. We all have little stories we tell ourselves and others, especially when we are asked why we did something. Lovers, for instance, are notorious for asking, “Why do you love me?” Often we are not really sure—we simply are drawn to the person; so much of what is important in relationships happens below the level of explicit awareness. However, we usually offer up some sort of account: “I like your eyes,” or “I like your enthusiasm.” We also each have a sort of personal story that we tell to ourselves and others—about how interesting, successful, ethical, honest, etc., we are. Are these phenomena at the normal end of a continuum, with confabulating people with neurological disorders at the other? Philosopher Daniel Dennett argued that one sort of meaning we can give to the overworked term “the self” is that the self is the subject of a story we create and tell to others about who we are: “Our fundamental tactic of self-protection, self-control, and self-definition is not spinning webs, but telling stories, and more particularly concocting and controlling the story we tell others—and ourselves—about who we are” (Dennett 1991, 418).

These stories have a unifying function according to Dennett: “These strings or streams of narrative issue forth *as if* from a single source—not just in the obvious physical sense of flowing from just one mouth, or one pencil or pen, but in a more subtle sense: their effect on any audience is to encourage them to (try to) posit a unified agent whose words they are, about whom they are: in short, to posit a center of narrative gravity” (1991, 418). Perhaps confabulation arises in part from this natural inclination toward telling stories about ourselves. Confabulation may also share with Dennett’s storytelling an unintentional quality: “And just as spiders don’t have to think, consciously and deliberately, about how to spin their webs, and just as beavers, unlike professional human engineers, do not consciously and deliberately plan the structures they build, we (unlike *professional* human storytellers) do not consciously and deliberately figure out what narratives to tell, and how to tell them” (1991, 418). In this Dennettian conception, confabulation is also a social phenomenon, but one that is more directly in the service of the individual than those mentioned earlier about the need for leaders to be confident. We will examine this view of confabulation in greater detail in chapter 7.

In an alternative conception, confabulation reflects the need that people have to assess their current situation and act on it quickly, without

pausing to consider all possibilities (Ramachandran and Blakeslee 1998). One of the brain's primary functions is to make sense of the world. Things and people must be categorized as this or that, as Tom or Mary, and decisions must be made about what to do with these things and people. Making categorizations always involves a risk of error, but doubt is a cognitive luxury and occurs only in highly developed nervous systems. One cannot focus too much on the details of what exactly the situation is, otherwise the most important issues are lost. Assuming it is a tiger rustling about in the bushes, what should I do about it? The brain has a proclivity to smooth over the rough edges and ignore certain details so as not to lose the big picture. Perhaps confabulation is a result of this sort of engineering strategy, but on a larger scale.

The brain likes to fill in gaps in the visual field, for instance. Each of your eyes has a blind spot, above and to the outside of that eye's focal center, where there are no rods or cones to receive light because the optic nerve exits the eyeball there. The reason you don't perceive a black spot in this portion of your visual field is that parts of the visual cortex fill in the blind spot, based on what is being perceived around its periphery (Ramachandran and Churchland 1994). For instance, if a solid patch of red is perceived around the periphery, these processes will fill in the blind spot with red, even though there might actually be a green spot there. Presumably the brain engages in these sorts of "coherencing" processes to facilitate its primary work: selection of actions from perceptions. Is the filling in of gaps in memory similar to filling in the blind spot, as Sully speculated long ago? "Just as the eye sees no gap in its field of vision corresponding to the 'blind spot' of the retina, but carries its impression over this area, so memory sees no lacuna in the past, but carries its image of conscious life over each of the forgotten spaces" (1881, 282). Does confabulation then belong with filling in as one of these "coherencing" or smoothing processes?

Aside from the questions about the causes and implications of confabulation for our understanding of the mind, there are also practical concerns that motivate our looking into the problem. In patients with neurological disease, confabulation is a barrier to the rehabilitation processes necessary for recovery, since they will not try to improve their memories, or their range of motion, until they acknowledge that a problem exists. The rest of us also need to know if we are prone to a problem that causes us to chronically misrepresent the world, and misunderstand and misrepresent to others how much we really know about ourselves or why we do things.

### **The Etymology of "Confabulation"**

According to the *Oxford English Dictionary* (OED 1971), "confabulation" is descended from the Latin term *confabulari*, which is constructed from the

roots *con*, meaning together, and *fabulari*, which means to talk or chat, so that the original meaning of “to confabulate,” is to talk familiarly with someone. Edgeworth wrote in 1801, for example, “His lordship was engaged in confabulation with his groom.” Confabulations are “talkings,” as Kempis translated in 1450: “consolacions are not as mannes talkinges or confabulacions.” The OED also notes, however, that *fabula* means tale, and evolved into the English word “fable.” So that “to confabulate with” another person is “to fable with” that person, as it were.

At the turn of the century, neurologists began applying a different sense of confabulation to some of their patients, beginning with those exhibiting what later came to be called Korsakoff’s amnesia, as in the example at the beginning of this chapter (Bonhoeffer 1904; Pick 1905; Wernicke 1906). The precise definition of the neurologist’s use of the word has been a subject of debate since it was introduced, however, owing to a lack of consensus as to what exactly the different essential features of confabulation are (Berlyne 1972; Whitlock 1981; Berrios 1998).

## 1.2 Confabulation Syndromes

It is widely accepted that confabulation comes in two forms, a milder version that includes the above examples, and a more severe and rare version in which the patient invents fantastic or absurd stories. The milder version was initially referred to as “momentary confabulation” by Bonhoeffer in his writings on patients with Korsakoff’s syndrome (1901, 1904). He called these “confabulations of embarrassment,” and speculated that they are created to fill in gaps in memory. Alternatively, what he referred to as “fantastic confabulation” overlaps heavily with things people who are under the influence of delusions say. Fantastic confabulations tend to have strange, “florid” (Kopelman 1991), or “extraordinary” (Stuss et al. 1978) content. Kopelman maintained, however, that the momentary-fantastic distinction “confounds a number of factors, which are not necessarily correlated, as the distinguishing features of the two types of confabulation; and it is wiser, perhaps, to focus attention upon one central feature by referring to ‘provoked’ and ‘spontaneous’ confabulation” (1987, 1482).

Provoked confabulation is produced in response to a question, whereas spontaneous confabulators produce their confabulations without being asked. Kopelman’s provoked–spontaneous dichotomy has gained some acceptance, and a body of recent research makes use of those concepts, for instance, in discussing whether they involve separate sites of damage (see Fischer et al. 1995; Schnider et al. 1996). An additional related question addresses whether the two types are discrete or merely represent gradations of severity. DeLuca and Cicerone (1991) resisted Berlyne’s claim

that “fantastic confabulation seems to be a distinct entity having nothing in common with momentary confabulation” (Berlyne 1972, 33), holding that the two types exist on a continuum. Consistent with this, there are reports of patients in whom spontaneous confabulation became provoked confabulation as the patients improved (Kapur and Coughlan 1980).

We will focus primarily on the provoked or milder form of confabulation, for a number of reasons. What is fascinating about provoked confabulation is that it occurs in people who are fully in possession of most of their cognitive faculties, and able to respond correctly to all sorts of requests and questions. Spontaneous confabulations are irrational stories that presumably result from delusions, and are seen primarily in schizophrenics. The value of provoked confabulation is that it shows the cognitive apparatus malfunctioning in an otherwise sound mind. It also promises to reveal valuable insights about the functioning of the normal cognitive system. The cognitive system of the schizophrenic confabulator is so severely broken that it is much more difficult to glean insights from it.

Before we take a closer look at the definition of confabulation, we need to look at some of the raw data. What follows is a survey of the different neurological syndromes known to produce confabulation, with several examples both in the clinic and in everyday settings.

Confabulation syndromes seen in the clinic include:

- Korsakoff's syndrome
- Aneurysm of the anterior communicating artery (ACoA)
- Split-brain syndrome
- Anosognosia for hemiplegia
- Anton's syndrome
- Capgras' syndrome
- Alzheimer's disease
- Schizophrenia

Confabulation has also been reported in normal people in certain circumstances, including young children, subjects of hypnosis, people asked to justify opinions, and people reporting mental states.

### **Confabulation in Clinical Settings**

Korsakoff's syndrome is a form of amnesia, most often caused by a lifetime of heavy drinking (Korsakoff 1889; Kopelman 1987). The memory deficit typically affects autobiographical memory (our memory of what happened to us) most severely, but it can also affect semantic memory (our impersonal knowledge of concepts and facts). The most frequent sites of damage are the mamillary bodies and the dorsomedial nuclei of the thalamus. Korsakoff's amnesia is severe enough that the patient will typically have no

memory at all of the events of the preceding day. However, when asked what he or she did yesterday, the Korsakoff's patient will often produce a detailed description of plausible (or not so plausible)-sounding events, all of it either entirely made up on the spot, or traceable to some actual but much older memory.

Berlyne gave the following example of confabulation from a patient with Korsakoff's syndrome: "L.M. was a 46-year old chronic alcoholic. He was admitted having been found at home in a very neglected state. He was euphoric, and showed devastation of recent memory. He said that he had been brought to the hospital by two Sick Berth Attendants and a Petty officer and that he was serving at Gosport on M.T.B.'s. (He had in fact been a Chief Petty Officer on motor torpedo boats during the war.) He said that the war was on and the invasion of Europe was imminent, yet he could recollect both VE and VJ day. He gave the date correctly; when told this would mean that the war had been going on for 20 years he was unperturbed" (Berlyne 1972, 32). Confabulatory patients often contradict themselves and, as with this man, show no interest in reconciling the contradictory claims.

Aneurysms of the anterior communicating artery, which distributes blood to a number of structures in the front of the brain, frequently produce confabulation (DeLuca and Diamond 1993; Fischer et al. 1995). They also produce a dense amnesia similar to that in Korsakoff's syndrome, leading some authors to speculate that confabulation is a deficit in monitoring information about "source memory"—the ability to place where in time and space a remembered event occurred, since the two syndromes share this feature.

The following is a description by Johnson and co-workers of a confabulating patient with an anterior communicating artery aneurysm: "G.S. had a number of erroneous ideas about his personal life to which he clung despite attempts to dissuade him. His fabrications were generally plausible and many involved autobiographical events that were embellished. A particularly salient confabulation was G.S.'s account of the origin of his medical condition; he believed that he had fallen and hit his head while standing outside talking to a friend, when in fact his aneurysm ruptured following an argument with his daughter" (Johnson et al. 1997, 192).

Aside from amnesia and confabulation, the third feature of ACoA syndrome is described as "personality change" (DeLuca 1993), and as with anosognosia, ACoA syndrome is accompanied by unawareness, in this case of the memory problem. Chapter 3 contains a detailed examination of this syndrome and its role as the primary piece of evidence for frontal theories of confabulation. It also contains a detailed inquiry into the locus of the lesion in Korsakoff's syndrome and any overlap it might have with lesion sites in ACoA syndrome.

Split-brain patients are epileptics who have had their corpus callosum surgically removed to lessen the severity and frequency of their seizures. In early testing of these patients in the 1960s, techniques were developed to send visual input to only one hemisphere or the other. Only the left hemisphere was thought able to give verbal responses, but it was found that the right hemisphere could understand pictorial and simple linguistic input, and could respond by pointing to pictures with the left hand [each hemisphere has control over the arm on the opposite side (Gazzaniga 1995a)]. When patients were asked about the activities of the left hand, though, the left hemisphere would answer as if *it* had been controlling the left hand, whereas, as a result of the operation, the left hemisphere had no idea why the left hand was doing what it was.

Several times during testing of the split-brain patients the left hemisphere produced reasonable-sounding but completely false explanations of the left hand's pointing behavior. In one study, a picture of a snow scene was lateralized to the right hemisphere of patient P.S., while a picture of a chicken claw was displayed to his left hemisphere. Then an array of possible matching pictures was shown to each hemisphere. P.S. responded correctly by pointing at a snow shovel with his left hand and at a chicken with his right hand. But when he was asked why he had chosen these items, he responded, "Oh, that's simple. The chicken claw goes with the chicken, and you need a shovel to clean out the chicken shed" (Gazzaniga 1995a, 225). In another study, a picture of a naked woman was shown only to the right hemisphere, using a tachistoscope (Gazzaniga and LeDoux 1978, 154). When the patient was asked why she was laughing, the left hemisphere responded, "That's a funny machine."

Anosognosia is exhibited by people with many types of neurological disorders, but it occurs most frequently after stroke damage to the inferior parietal cortex of the right hemisphere (Bisiach and Geminiani 1991; Heilman et al. 1991). Damage here can produce paralysis or great weakness of the left arm or of the entire left side of the body. This paralysis can be accompanied by neglect, a condition in which the patient ignores the left side of the body and its surrounding space. A patient with neglect typically will not eat food on the left side of the plate or wash the left side of the body, and will not notice people standing quietly on her left. Some patients with this left-side paralysis and neglect also exhibit anosognosia for several days after a stroke. Approached on her right side as she lies in bed and asked whether she can use her left arm, such a patient will answer matter-of-factly that she can. When the neurologist tells the patient to touch his nose with her left arm, the patient will try in vain to reach it, or will occasionally reach out with her right arm instead. But often, she will produce a confabulation, saying something such as, "I have never been very ambidextrous," or "These medical students have been probing me all day and I'm sick of it. I don't want to use my left arm" (Ramachandran 1996b, 125).

When asked whether she reached successfully, the patient who tried to reach will often say that she did, and a large percentage of these patients will claim that they saw their hand touch the doctor's nose.

Neglect and denial can also be accompanied by *asomatognosia*, in which the patient denies that the paralyzed or greatly weakened arm even belongs to him or her. Sandifer (1946, 122–123) reported the following dialogue between an examining physician and an anosognosic patient:

Examiner (holding up patient's left hand): "Is this your hand?"

Patient: "Not mine, doctor."

Examiner: "Whose hand is it then?"

Patient: "I suppose it's yours, doctor."

Examiner: "No, it's not; look at it carefully."

Patient: "It is not mine, doctor."

Examiner: "Yes it is, look at that ring; whose is it?"

Patient: "That's my ring; you've got my ring, doctor."

Examiner: "Look at it—it is your hand."

Patient: "Oh, no, doctor."

Examiner: "Where is your left hand then?"

Patient: "Somewhere here, I think." (Making groping movements toward her left shoulder.)

The patients give no sign that they are aware of what they are doing; apparently they are not lying, and genuinely believe their confabulations. They do not give any outward signs of lying, and their demeanor while confabulating was described by Kraepelin (1910) as "rocklike certitude." In one experiment that affirms the sincerity of confabulators, anosognosics with left-side paralysis were given the choice of performing a two-handed task (tying a shoe) for a reward of \$10, or a one-handed task (screwing a light bulb into a socket) for \$5. The patients uniformly selected, then failed at the two-handed task. In contrast, those who had left-side paralysis caused by a right hemisphere stroke but no anosognosia systematically chose the one-handed task (Ramachandran 1995). In the clinic, anosognosics are often observed trying to use their paralyzed left arms or legs.

The fact that the overwhelming majority of neglect patients ignore the left side of personal space as a result of right hemisphere damage seems to lend support to the claim that the causes of confabulation have an important lateral component. Confabulating neglect patients share with split-brain patients the feature of having a left hemisphere cut off from the information it needs to answer the doctor's questions. Chapter 6 examines these connections after giving a thorough description of what is known about this type of anosognosia.

Another type of anosognosic patient denies that he or she is blind. Known as Anton's syndrome, this is a rare condition that is typically

traceable to bilateral damage to the occipital lobes, areas at the back of the brain specialized for visual processing (Anton 1899; Swartz and Brust 1984), coupled with frontal damage. The occipital damage can cause a condition known as cortical blindness—blindness that is due to cortical damage rather than damage to the eyes or optic nerves. A small percentage of patients with cortical blindness also exhibit Anton’s syndrome; they tend to have either diffuse damage caused by dementia, or circumscribed frontal lobe lesions (McDaniel and McDaniel 1991). Those with Anton’s syndrome can be quite rational in general, until one asks them to describe what they see. They typically produce a description that is logical or plausible, but false. For instance, if asked to describe what their doctor is wearing, they will provide a full description of a generic doctor. When confronted with the falsity of the description, the patients make excuses similar to those of patients who deny paralysis. Benson’s patient with Anton’s syndrome “adamantly denied any visual problems, often complaining that the light was poor in the room, that he did not have his best pair of glasses with him, or that it was nighttime” (1994, 87).

Capgras’ syndrome is a rare condition in which a patient claims that people close to him or her, typically parents, spouses, or children, have been replaced by impostors (Capgras and Reboul-Lachaux 1923). The locus of the lesion is unknown, but a consensus is building that it involves a temporal lobe lesion in concert with a frontal lesion (e.g., Signer 1994). The most popular current hypothesis about why the impostor delusion is formed is that the patient is missing a feeling of emotional arousal that the sight of a significant person normally produces, and that the impostor claim is a confabulation created to explain why the person feels different to the patient. V. S. Ramachandran and I asked a patient with Capgras’ syndrome point-blank why someone would pretend to be his father. His reply was: “That is what is so strange, Doctor—why should anyone want to pretend to be my father? Maybe my father employed him to take care of me—paid him some money so he could pay my bills” (Hirstein and Ramachandran 1997, 438). This patient, D.S., also had the same illusion when presented with photographs of himself, as shown in the following interchange (1997):

Examiner (Pointing to photograph of D.S. taken two years ago when he had a moustache): “Whose picture is this?”

Patient: “That is another D.S. who looks identical to me but he isn’t me—he has a moustache.”

Capgras’ syndrome is different from many other confabulation syndromes in that the confabulation posits something different from the status quo, compared with the denials of anosognosics, which seem to be designed to

affirm the patient's old body image. The behavior of Capgras' patients tends to be consistent with their impostor claims; there are reports of patients killing the "impostors," for instance.

Alzheimer's disease involves the loss of neurons in several different brain areas, and manifests itself in a progressive loss of abilities in several different cognitive tasks (Wells and Whitehouse 1996). It is the most common type of what neurologists call dementias. Confabulations produced by these patients tend to be similar to those produced by people with Korsakoff's syndrome (Kopelman 1987, 1991). The conventional wisdom is that confabulation occurs in Alzheimer's disease when the diffuse cortical atrophy reaches a point at which the frontal lesions necessary to produce confabulation occur (see Kern et al. 1992), perhaps overlapping with the lesions found in Korsakoff's or in aneurysms of the anterior communicating artery.

Schizophrenia is a serious psychological disorder that involves large-scale disruptions in perception, emotions, thinking, and behavior (Jeste et al. 1996). Unlike some of the other syndromes we have discussed, there is no widely agreed-upon theory as to the locus of damage in the schizophrenic brain. Nathaniel-James and Frith (1996) first broached the idea that confabulation is present in schizophrenics (see also Kramer et al. 1998). They read narratives adapted from Aesop's fables to schizophrenic patients, then immediately afterward asked the subjects to recall as much of the story as they could. When the subjects inserted features that were not present in the story, these were counted as confabulations. Nathaniel-James and Frith suggested that schizophrenics share with frontal damage patients the inability to suppress inappropriate responses, something that links confabulation with the phenomenon of disinhibition (there is more on this in chapter 4).

### **Confabulation in Normal People**

Young children sometimes confabulate when asked to recall events. Ackil and Zaragoza (1998) showed first-graders a segment of a film depicting a boy and his experiences at summer camp. Afterward the children were asked questions about it, including questions about events that did not happen in the film. One such question was, "What did the boy say Sullivan had stolen?" when in fact no thefts had taken place in the film. The children were pressed to give some sort of answer, and the experimenters often suggested an answer. When the children were interviewed a week later, the false events as well as the suggested answers had been incorporated into their recollections of the movie. These false memories are discussed further in chapter 3.

Subjects of hypnosis may confabulate when they are asked to recall information associated with crimes (Dywan 1995, 1998), causing

researchers to warn criminologists about the dangers of obtaining information from hypnotized subjects. There are also anecdotal reports of hypnotized subjects confabulating when asked why they did something in accord with their hypnotic suggestion. For instance, a stage hypnotist gives his subject the suggestion that he will wave his hands whenever he hears the word "money." When asked later why he is waving his hands, the subject replies "Oh, I just felt like stretching."

In addition to the experiment described earlier, in which shoppers were asked their preferences about nylon stockings, Nisbett and colleagues conducted several other experiments that also seem to show confabulation by normal people. In one study, subjects with insomnia were given a placebo pill and were told it would produce rapid heart rate, breathing irregularities, bodily warmth, and alertness—all normal symptoms of insomnia (Storms and Nisbett 1970). The experimenters' idea was that knowledge of having taken the pill would cause the subjects to fall asleep earlier because they would be able to attribute their symptoms to the pill, rather than to whatever emotional turmoil in their lives was actually producing the insomnia. This is exactly what happened; the subjects reported going to sleep more quickly. However, when asked why they fell asleep earlier, they seemed to confabulate: "Arousal subjects typically replied that they usually found it easier to get to sleep later in the week, or that they had taken an exam that had worried them but had done well on it and could now relax, or that problems with a roommate or girlfriend seemed on their way to resolution" (Nisbett and Wilson 1977, 238).

Philosophers have found the concept of confabulation useful in describing people's reports of what went on in their minds during certain tasks such as solving puzzles (see also Nisbett and Wilson 1977). Philosophers of a behaviorist bent find the idea of confabulation amenable to their contention that we do not have reliable access to what goes on in our minds; that is, that introspection is not to be understood on the model of seeing, and that reports of introspections are not similar to reports of seen events. Dennett says, "there are circumstances in which people are just wrong about what they are doing and how they are doing it. It is not that they lie in the experimental situation but that they confabulate; they fill in gaps, guess, speculate, mistake theorizing for observing. . . . They don't have any way of 'seeing' (with an inner eye, presumably) the processes that govern their assertions, but that doesn't stop them from having heartfelt opinions to express" (1991, 94).

These normal people seem similar to the split-brain patients described earlier, who confabulated about why their left hands performed certain actions. In chapter 7 we examine the idea that certain reports of mental states are confabulations.

### 1.3 Features of Confabulation

Two very different sorts of activities might equally well be described as defining confabulation. The first activity has as its goal defining the word based on how people use it; the second focuses on determining what confabulation itself is. To put it another way, the first activity describes what the word *means* to people who use it, and the second focuses on what the word *refers* to.

The first activity involves discovering what criteria or conditions people normally apply when they use a word. The characteristics that we apply for “gold,” for instance, include a yellowish color and a somewhat soft and heavy consistency compared with other metals. What actually makes a metal gold, however, is that it is made up of a certain type of atom; it has an atomic weight of 79. Some substances may look and feel like gold, such as iron pyrite (fool’s gold), but are not gold because they lack the atomic and molecular structure of true gold (Putnam 1971; Kripke 1977).

Because of the distinction between the meaning and referent of a word, people are sometimes wrong about what they are referring to. People thought that “jade,” for instance, referred to a single type of green stone, which makes elegant jewelry. It turns out that they were wrong. There are two chemically different types of stone to which this term is applied, one of which is much more attractive (jadeite) and valuable than the other (nephrite). Closer to our topic, “schizophrenia,” for example, may turn out to refer to more than one type of brain disorder (Jeste et al. 1996).

The second activity in defining confabulation involves delineating the actual phenomena that people refer to when they use the word, even if they are wrong in their beliefs about them. Each of the seven criteria discussed in the following pages can initially be seen either as part of the meaning of the term “confabulation,” or as designating its actual features. As with jade and gold, these two aspects can be teased apart, which we will have to do on certain occasions. We might agree that a particular criterion is currently part of the meaning of confabulation, but find on investigation that it is not actually a feature of the phenomenon once we understand what confabulation actually is. These criteria and their proposed features are discussed next.

#### Criteria for Confabulation

**1. Does the Patient Intend to Deceive?** The orthodox position is that the patient has no intent to deceive. Ramachandran’s result, in which patients who denied their paralysis attempted two-handed tasks, is one piece of evidence in support of this claim. Confabulation is not lying, which involves clear intent to deceive.

I lie to you when (and only when)

1. I claim  $p$  to you.
2.  $p$  is false.
3. I believe that  $p$  is false.
4. I intend to cause you to believe  $p$  is true by claiming that  $p$  is true.

Confabulators do not satisfy the third condition since they seem to steadfastly believe what they say. The following dialogue, from DeLuca (2001, 121), shows a man with an anterior communicating artery aneurysm who is sincere about what he claims:

Doctor: You indicated last night you were working on a number of projects at home. . . . What would you say if I told you you were actually here in the hospital last night?

Patient: I'd be surprised, because my experience, what I learn from my eyes and ears tells me differently. . . . I'd want some evidence. I'd want some indication that you knew about my private world before I gave any cognizance.

Doctor: Would you believe me?

Patient: Not out of the blue, especially since we haven't even met (an illustration of the patient's amnesia).

Doctor: What if your wife was here and she agreed with me, what would you think at that point?

Patient: I'd continue to resist, but it would become more difficult.

## 2. Does the Patient Have Some Motive Behind His or Her Response?

This criterion is of course related to the first one about intent to deceive; lying typically functions in the service of some motive. The most obvious motive in the case of confabulation would be a desire to cover up one's deficit. Some older textbooks combine the questions of motive and deception in a way that looks as if the authors are claiming that deception is intentional. Freedman et al. said that the patient recovering from amnesia after head injury "usually has a tendency to confabulate in order to cover his recent memory defect" (1975, 1428). Whitlock (1981) took these authors to task for their apparent attribution of intention to deceive, but the authors may not actually have been guilty of this. They were stumbling over a problem philosophers have encountered in their investigation of the problem of self-deception—how to describe behavior that seems to fall between the intentional-unintentional distinction. We will pursue this question in detail in chapters 8 and 9.

Another motive sometimes cited is the desire to avoid something known as the "catastrophic reaction," where the patient comes to the horrible realization that he or she is paralyzed, has lost her memory, etc. and becomes depressed (Gainotti 1975; Zangwill 1953). A third possible motive

that may also be at work in normal cases of confabulation is simply the desire to avoid saying, "I don't know," especially when the provoking question touches on something people are normally expected to know.

**3. Must a Defective Memory Be Involved?** Because it was first encountered in memory disorders, confabulation is traditionally defined in such a way that it must be accompanied by a memory deficit. Berlyne's classic definition of confabulation is that it is "a falsification of memory occurring in clear consciousness in association with an organically derived amnesia" (1972, 38). Similarly, Mercer et al. stated that "a necessary (though not sufficient) prerequisite for confabulation is impaired memory function" (1977, 433). Patients with Capgras' syndrome, however, do not have an obvious memory deficit. Nathaniel-James and Frith (1996) also argued that their schizophrenic patients exhibit confabulation in the absence of a memory deficit (see also Weinstein et al. 1956; Weinstein 1996). Similarly, confabulation in split-brain patients and in anosognosics is not accompanied by any obvious memory deficit. Even in the case of the memory confabulation syndromes, such as Korsakoff's, we will see that there are several indications that amnesia and confabulation are caused by damage to different brain areas.

**4. Must the Confabulation Be in Response to a Question or Request?** The traditional approach is of course covered by the distinction between *spontaneous* and *provoked* confabulation. This difference is important, since the questioning of the examiner sets up a special context in which an authority figure is soliciting information, which is not present in the case of spontaneous confabulation. Several outcomes seem available here. It might turn out that spontaneous confabulation is not actually confabulation, but simply the expression of a delusion. Alternatively, it might turn out that there are two types of confabulation, and the spontaneous-provoked distinction does not draw the correct boundary between the two types.

**5. Does the Confabulation Fill a Gap?** According to this criterion, confabulations fill in gaps at a certain level in the cognitive system. Perhaps this is because confabulation is another example of a tendency exhibited at many levels of brain function, to produce complete, coherent representations of the world. There are, however, several problems with Korsakoff's idea that confabulation fills a gap in the patient's memory. For one thing, a patient with Korsakoff's syndrome does not merely have a gap in episodic memory about recent activities; there is nothing there at all. The "gap" is not properly filled either, since it is filled with a (probably) false claim, with mud rather than mortar. It might seem *to the patient*, however, that he has

merely a gap in his memory, and that confabulation fills that gap; but again, this may be making it all too intentional.

**6. Are Confabulations Necessarily in Linguistic Form?** Several researchers have categorized nonlinguistic responses as confabulatory. Lu et al. (1997) had patients point to fabric samples with one hand to indicate which texture of fabric they had been stimulated with on the other hand. The patients also had the option of pointing to a question mark in trials in which they had not been stimulated, a nonlinguistic version of answering “I don’t know.” The authors operationally defined confabulation as failure to point to a question mark. Bender et al. (1916) applied the term “confabulation” to the behavior of patients when they produced meaningless drawings as if they were familiar designs. Similarly, Joslyn et al. (1978) had patients reproduce from memory certain drawings they had seen, and described cases in which the patients added extra features that were not actually present as confabulations (see also the work of Kern et al. [1992] on Alzheimer’s disease, and Chatterjee [1995] on neglect).

These uses of confabulation seem to conflict with the idea that confabulators tell stories, which are usually false. Pointing at a piece of cloth, or answering “yes” rather than “I don’t know” to a question do not seem to be confabulations according to this view. Those researchers seemed to be conceiving of confabulation as a broader epistemic phenomenon, rather than as a narrower, purely linguistic one. In the epistemic view, a confabulation is a poorly grounded claim, and a confabulatory person tends to make epistemically ill-grounded claims. Similarly, “to confabulate,” means something like “to confidently claim something one has a poor epistemic warrant for.”

For confabulations that are linguistic, a further question is whether they must be internally consistent. Weinstein (1996) argued that confabulations should be coherent, as did Talland (1961). Contrary to this, Moscovitch (1995), stated that they “need not be coherent and internally consistent”; there are several examples of confabulation, even of the provoked variety, in the literature that contain contradictions. Another semantic question about confabulations that are responses to questions is whether they all must be false. A confabulatory patient may on occasion say something true, perhaps by accident.

**7. Are Confabulations the Result of Delusions?** That is, what is the relation between confabulation and delusion? It would be wrong to classify even spontaneous confabulation as a type of delusion, since delusions are, minimally, false or ill-grounded beliefs (but see Stephens and Graham 2004), whereas confabulations are false (or ill-grounded) claims. (Similar questions arise about the wisdom of making falsity part of the definition of

delusion, as arise in the case of confabulation; see Fulford 1989.) A delusion might give rise to a confabulation, however, which seems to be what happens in the case of Capgras' delusion. One way in which confabulation was differentiated from the expression of a delusion in the past was by the claim that confabulation necessarily involves a memory deficit (Berrios 1998). However, with the addition of several syndromes that do not involve memory deficits to the list of confabulation syndromes, this criterion may have lost its usefulness. Delusions also tend to be long lasting; they are described by the American Psychiatric Association (1994, 765) as "firmly sustained in spite of what almost everyone else believes and in spite of what constitutes incontrovertible and obvious proof or evidence to the contrary." Many confabulations, on the other hand, are quickly forgotten after they are communicated. Confabulators do tend to resist changing their claims in the face of contrary evidence, but not with the tenacity of the deluded.

In chapter 8 after we have surveyed the data on confabulation, we will construct a definition of confabulation. There we will examine several of these possible criteria for inclusion in the definition.

#### 1.4 Three Concepts of Confabulation

Once the different possible features of confabulation are described, it becomes clear that there are competing concepts of confabulation, each of which makes different sets of the features described here essential to the definition. We might think of a concept as a bundle of such features, with the more essential features at the center and the less essential ones at the periphery. I intend *concept* to have a certain neutrality or hypothetical nature. A concept can crystallize into a definition, which has happened in the case of the mnemonic concept. Two other concepts seem to be at work in the minds of people using "confabulation"—the linguistic and the epistemic—but they have not yet been formally defined.

##### **Mnemonic Concept**

According to the classic concept, confabulations are stories produced to cover gaps in memory. Memory and gap-filling features are essential in this definition. Since its introduction at the beginning of the twentieth century, however, the concept has been increasingly threatened by applications of confabulation to patients who have no obvious memory problem. Because of its reliance on the notion of gap filling, the mnemonic concept also must address the question of whether the confabulator does this knowingly, with intent to deceive. Most people employing the classic concept agree that confabulators do not deliberately attempt to deceive; thus Moscovitch's (1989) description of confabulation as "honest lying."

### Linguistic Concept

Here, confabulations are false stories or sentences. In this concept, confabulation is a symptom shared by a set of syndromes with diverse physical bases, in the way that high blood pressure is the name of a symptom with many different causes. The noun *confabulation* is defined first as a false narrative, and the verb *to confabulate* is the act of producing a confabulation, in the noun sense. This concept is truer to the linguistic meaning in the etymology of confabulation as an act of conversing, or talking with someone. Proponents of the linguistic concept might trace the roots of confabulation to the penchant we humans have for telling and listening to stories. Since this concept is neutral about the question of whether the person intentionally produced a false narrative, it can avoid the difficult question of intention that other concepts are saddled with.

One objection to the linguistic concept is that it causes us to mistakenly lump lying and confabulation together because both involve the production of false narratives. Its emphasis on the language itself also poses a special problem for this concept concerning the truth of confabulations. Are confabulations false by definition? Some of the memory claims of patients with Korsakoff's syndrome are true but misplaced in time. Similarly, even a highly confabulatory patient will occasionally make true claims. The obvious way around this, to note that a patient might happen to say something true but that it would be a matter of luck, pushes the linguistic concept closer to the epistemic concept. On the other hand, if different syndromes that result in confabulation are not found to share any significant physical basis, this will lend support to the linguistic concept; all they have in common is a symptom: the production of false narratives.

### Epistemic Concept

In this concept, a confabulation is a certain type of epistemically ill-grounded claim that the confabulator does not know is ill-grounded. The claims need not be made in full natural language sentences; they may consist of drawings, pointing to a picture, or simple "yes," and "no" answers to questions. Dennett's use of confabulation, quoted earlier, "It is not that they lie in the experimental situation but that they confabulate; they fill in gaps, guess, speculate, mistake theorizing for observing," (1991, 94) seems to have the epistemic concept behind it. Studies cited earlier that counted pointing to a cloth sample or producing certain drawings as confabulation also may be employing an epistemic concept. Both activities involve making a claim, the claim that I was stimulated with *this* type of texture, or *this* is the drawing I saw.

One way to delineate the epistemic concept further would be in terms of certain malfunctions of normal belief forming and expressing processes (Goldman 1986). The malfunctioning process in the case of con-

fabulation may be one that allows people to attach doubt to ideas; confabulatory people do not experience doubt about their claims and ideas, whereas a normal person would. The claims are epistemically ill-grounded because they have not passed a review process that can result in doubt about them. That process does not function correctly because of brain damage. This is one way to explain the connection between confabulation and the disinhibition that frequently accompanies it; the doubt normal people feel when they consider certain ideas is caused by the same process that normally inhibits inappropriate responses.

The groundedness of our claims seems to come in continuously variable degrees, over which we normally have some voluntary control. We can loosen or tighten our epistemic criteria, depending on the situation. When people give legal testimony, for instance, they are able to apply abnormally high standards to what they say, much higher than when they are speaking casually with friends. We also take this sort of care in some everyday situations: a man warns his friend before they meet with someone, "Be careful what you say to him." Have confabulating patients perhaps lost this ability to regulate the level of certainty each context requires?

By understanding confabulation as arising from malfunctioning epistemic processes, this concept is committed to going beyond the expression of confabulation to the processes behind it. One objection to this concept is that it applies confabulation too broadly and hence reduces the likelihood of our finding a single neural basis for it. The way to respond is presumably to argue in favor of a malfunction of a general, high-level brain process, one that is involved in all the syndromes of confabulation.

### 1.5 Mirror-Image Syndromes

Sometimes one can understand something better by asking what the opposite of that thing is. Constructing a system of opposites also helps to provide a large conceptual space in which to place confabulation. The existence of such opposite or mirror-image syndromes also helps to dispel any notion that confabulation is somehow a universal feature of any sort of serious brain damage. Of course, there is no such thing as *the* opposite of something; rather, something is the opposite of another thing in a certain respect. One feature of confabulation is the confabulator's claim that he has certain abilities that he in fact lacks, such as the ability to remember, to move his arm, or, in the case of split-brain patients, to explain why his left arm did something. In contrast with this, there is a class of neurological syndromes in which patients *do* possess certain abilities, but claim that they do not. *Blindsight* patients, for example, have a large blind spot in their visual fields that is due to an injury, and will claim that they cannot see anything at all there (Weiskrantz 1986). If they are asked which way a

beam of light moved around in the blind area, they are able to answer correctly because other, nonconscious streams of visual processing are intact. In a case even more closely related to confabulations about vision, Hartmann et al. (1991) described a patient with cortical blindness who suffered from what they called “inverse Anton’s syndrome.” The man claimed to have no ability to see at all, yet testing revealed that he had a small area of the visual field in which his visual abilities were well preserved, including the ability to name objects, colors, and famous faces, as well as to read words.

Similarly, some patients with *prosopagnosia* (*face agnosia*), who claim to be unable to visually recognize familiar people, show unconscious autonomic reactions (as measured by activity in the sweat glands of their hands) to photographs of people they know (Bauer 1984; Tranel and Damasio 1985), just as normal people do. Bauer (1984, 1986) and Ellis and Young (1990) hypothesized that prosopagnosics suffer from a syndrome that is the mirror image of Capgras’ syndrome, suggesting that Capgras’ patients recognize relatives (albeit as impostors) but fail to register the normal skin-conductance response to the sight of them, a prediction that has been confirmed experimentally (Ellis et al. 1997; Hirstein and Ramachandran 1997). We examine the distinction between prosopagnosia and Capgras’ syndrome in greater detail in chapter 5.

We noted that another feature of confabulation seems to be an inability to doubt what one says; it involves a sort of pathological certainty. The opposite of this would be pathological doubt. Two sorts of conditions are possible opposites in this regard: paranoia and obsessive-compulsive disorder (OCD). Paranoia may not be a good candidate, since it has been observed to coexist with confabulation (Berlyne 1972; Weinstein 1996), and one might contend that it involves more suspicion than doubt. Obsessive-compulsive disorder, however, may be a good candidate. It can be interpreted as pathological doubt—doubt that one’s hands are clean, or that all the doors are locked, or that the stove has been turned off, for example (Schwartz 1998; Saxena et al. 1998). A possible explanation of the relation between confabulation and OCD is that the process that produces a feeling of doubt and is hyperfunctioning in OCD is completely broken in confabulatory patients (see chapter 4).

## 1.6 Conclusion: Setting the Problem of Confabulation

The problem of confabulation is one of giving a satisfactory explanation for what exactly the phenomenon is, and what it tells us about who we are, how we think, and how the brain works. An account of confabulation should be able to answer the following questions:

1. Do all behaviors referred to as confabulations involve the same brain malfunction?
  - a. What are the important features of the phenomenon (or phenomena if it is not unitary)?
  - b. What causes confabulation?
  - c. Why does confabulation appear in so many disorders?
  - d. Are there different subtypes of confabulation? What are they?
  - e. What is the connection between confabulation and denial of illness?
  - f. Why are some disabilities admitted and others denied?
  - g. How can confabulation be prevented or treated?
2. What does the existence of confabulation tell us about human nature?
  - a. Is confabulation only a pathological phenomenon, or is there a continuum, shading from pathological cases into normal behavior?
  - b. What is the connection between confabulation and the self?
  - c. What is the connection between confabulation and self-deception? Are statements made by self-deceived people confabulations?
  - d. Does confabulation have positive functions?

Obviously, those adhering to different concepts—mnemonic, linguistic, and epistemic—have different answers to many of these questions. A comparison of how the concepts handle the phenomena of confabulation will prove informative in what follows. But more important, the interplay between the concepts and the empirical evidence will, I hope, lead to our understanding of what this fascinating phenomenon is. In the ensuing chapters, I propose that the epistemic concept is preferable, both because it nicely captures a natural set of the phenomena of confabulation, but also because it provides a productive and substantially correct guide to investigating and understanding the brain malfunctions behind confabulation.