
This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: http://openaccess.city.ac.uk/14569/

Link to published version: http://dx.doi.org/10.1016/j.concog.2014.12.002

Copyright and reuse: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.
Remembering, imagining, false memories & personal meanings

Martin A. Conway a,*, Catherine Loveday b,1

a Centre for Memory & Law, Department of Psychology, City University London, United Kingdom
b Department of Psychology, University of Westminster, United Kingdom

Article info
Article history:
Received 26 July 2014
Revised 3 December 2014
Accepted 3 December 2014
Available online 12 January 2015

Keywords:
Autobiographical memory
Episodic memory
Episodic future thinking
Simulation
Imagining
Working self
Personal meanings

Abstract
The Self-Memory System encompasses the working self, autobiographical memory and episodic memory. Specific autobiographical memories are patterns of activation over knowledge structures in autobiographical and episodic memory brought about by the activating effect of cues. The working self can elaborate cues based on the knowledge they initially activate and so control the construction of memories of the past and the future. It is proposed that such construction takes place in the remembering–imagining system – a window of highly accessible recent memories and simulations of near future events. How this malfunctions in various disorders is considered as are the implication of what we term the modern view of human memory for notions of memory accuracy. We show how all memories are to some degree false and that the main role of memories lies in generating personal meanings.

© 2014 Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

1. Introduction

The main contention of this paper is that when people remember they imagine and when they imagine they use memory. Imagining involves ‘working with memory’ (Moscovitch, 1992). Because of the intrinsic relatedness of memory and imagination we refer to what we term the remembering–imagining system, RIS, (Conway & Loveday, 2015). The RIS is described further below. The constructive nature of autobiographical memory and autobiographical remembering are considered first.

2. Autobiographical memory

Autobiographical memory (AM) is a complex cognitive system mediated by neural networks distributed through large areas of the neo-cortex and limbic system (see Cabeza & St. Jacques, 2007). Indeed, recent neuroimaging studies have found few differences between remembering, imagining the future, and what is sometimes termed ‘the default network’, all of which appear to share the same extensive distribution of interlocking neural networks (see Schacter, Chamberlain, Gaesser, & Gerlach, 2012 for a recent review). Autobiographical memory contains autobiographical knowledge, e.g. personal factual knowledge and cultural knowledge, such as the history of our times. It also contains episodic memories,
e.g. fragmentary knowledge derived from experience (Conway, 2009). As such it forms a major part of the self (Conway, 2005; Conway & Pleydell-Pearce, 2000; Conway, Singer, & Tagini, 2004). Fig. 1 depicts partonomic knowledge structures in AM in which episodic memories are part of general events which in turn are part of lifetime periods which may themselves be part of broader themes such as work or relationship themes and the life story (Bluck & Habermas, 2001).

Different levels of the knowledge structures are accessed by cues and the lines in Fig. 1 connecting different levels depict the action of cues and should not be taken as some sort of direct connection. So for example, knowledge represented in a lifetime period such as ‘Working at university X’ could be used as a cue to access the general event, ‘Departmental talks’, which in turn contains knowledge that can access specific episodic memories. Thus, the entire complex knowledge base has patterns of activations arising and dissipating continuously in it as experience is represented in the mind and has the effect of activating associated long-term knowledge. The idea is that the AM system is labile and intrinsically responsive to cues. Occasionally, on a daily basis according to Berntsen and colleagues (Berntsen, Staugaard, & Sørensen, 2013), an AM may spontaneously come to mind. Such involuntary memories often appear to be the result of a specific cue, illustrating this cue-sensitivity. On other occasions control process elaborate a cue as a specific memory is sought for (Conway, 2005). This process of generating a memory takes time as a cue iterates through repeated cycles of elaboration and activation, until the sought-for memory has been constructed. In all cases the effect of a cue or set of cues is to form a stable pattern of activation within AM knowledge structures and it is that pattern of activation that is, temporarily, a memory. A specific AM will also always include the activation of episodic memories.

One interesting hypothesis here is that it is only when an episodic memory or set of episodic memories are activated that a constructed memory enters consciousness, i.e. the rememberer becomes consciously aware of the memory. A possibility that then arises is that activation in the autobiographical memory knowledge base can, and indeed does, occur non-consciously. Experiences of ‘involuntarily’ remembering may then simply reflect unawareness of memory processing that occurred prior to the ‘involuntary’ recall. Indeed, Schank (1982) catalogued many instances of what he term ‘reminders’ – memories unexpectedly coming to mind often because of an abstract relation to a current situation or to another memory. So, for example, the structure of an event might cue a memory, as in the case of a person who recalled that he could not get his hair cut as short as he wanted when in England cueing a memory of not being able to get a steak cooked as rare as he liked, cf. Schank (1982).

Reminders suggest that there maybe some non-conscious process that monitors memory for autobiographical knowledge that could help current problem solving. The non-conscious activation of memories may also underlie feelings such as déjà states, e.g. déjà vu, having seen before, and déjà vecu – having lived this moment before. We have proposed that in patients who often experience déjà states, particularly déjà vecu, their memory feelings may arise from memories that do not enter into consciousness but nevertheless are strongly enough activated to trigger recollective experience (Moulin, Conway, Thompson, James, & Jones, 2004). One of the problems with a memory system that has highly labile patterns of activation continually arising in response to continuously changing cues is that memories could, potentially, swamp consciousness. Thus, control process must modulate what memories become instantiated in consciousness at the same time as monitoring the cue-driven changing patterns of activation in long-term memory for information relevant to current goals. Finally, it seems to us highly possible that such non-conscious activation of autobiographical memory knowledge structures might, potentially, play a significant role in mediating a wide variety of social interactions, e.g. in influencing who we like, who we do not like, and who we are uninterested in, although this aspect of autobiographical memory has yet to be investigated.

Autobiographical memory is then a complex multilayered distributed knowledge base in which cues constantly cause patterns of activation, some of which may stabilize into memories. It seems that this AM network is never totally inactive although, of course, activation will vary in strength at different times and may be modulated by control processes as suggested above. For instance, neuroimaging studies have found a ‘default’ network that is active when attention is unfocused but which is powered down when attention is focused (Schacter et al., 2012). The default network takes in many of the networks that feature in autobiographical remembering and, importantly, many of the same networks are active when future experiences are imagined. Thus, during ‘day dreaming’ the autobiographical memory knowledge base is active. In addition during sleep it is now clear that a major component of the AM system, the medial temporal lobe memory system is also highly active (see Solms, 1997). The AM knowledge base is, then, never totally quiescent and regions of it are always active, even during sleep.

In general, however, it may be the case that control or executive processes exercise considerable control over which patterns of activation enter consciousness and also how patterns of activation are ‘shaped’ by cues. The latter by elaborating cues on the basis of initial cue activation and so directing a search. The former by denying patterns of activation access to consciousness, where they would disrupt current goal processing by turning attention to the pattern of activation in the AM knowledge base and away from other attention focused activities. Despite this inhibitory control the effect of cues is always to cause activation in the AM knowledge base. As proposed earlier it seems possible that such activation, that does not gain conscious representation, may nonetheless influence goal processing non-consciously. Whether this actually occurs is not known, but the Self–Memory System (SMS) model of AM put forward by Conway and Pleydell-Pearce (2000) allows for the possibility.
3. Autobiographical remembering

Fig. 2 shows a subsequent development of the SMS model that distinguishes between the working self, autobiographical memory and episodic memory. The working self consists of the conceptual self (Conway, Singer, et al., 2004) and the goal system. Memories are shaped in their construction by the working self and the working self determines what knowledge derived from experience becomes encoded into the AM knowledge base. Note that, as all experience is internal, encoding includes sensory-perceptual information, affect, thoughts, imaginings – it is a sort of derivation of the cognitive milieu over a limited time, perhaps from an event beginning to event ending (see Williams, Conway, & Baddeley, 2008). In this version of the SMS, AM consists of autobiographical knowledge. Autobiographical knowledge includes personal knowledge such as lifetime periods and general events (see Fig. 2) and also more generic and cultural knowledge.

Lastly, the notion of episodic memory is developed further in this model of the SMS (Conway, 2009). Simple episodic memories consist of fragments of knowledge derived from experience, often in the form of visual images, together with...
‘frame’, i.e. contextualizing conceptual knowledge. Simple episodic memories can be grouped together, on the basis of shared cues, to form complex episodic memories, as shown in Fig. 2. Patterns of activation over episodic and autobiographical memory constitute specific autobiographical memories. The whole SMS is self-constraining and the working self determines what memories can be accessed whereas autobiographical memory constrains what goals, plans, beliefs, etc. can be realistically held (see Singer & Conway, 2011, for a psychoanalytic extension of this model).

4. The remembering imagining system (RIS)

The RIS is a window of accessibility of memories of the recent past and the near future (Conway & Loveday, 2015). Recently formed episodic memories e.g. those formed today, are highly accessible, however, as the retention interval increases these become progressively less accessible. Similarly near future (imagined) events, e.g. those anticipated to occur tomorrow, are also highly accessible and become progressively less accessible as the expectation interval increases. Conway and Loveday (2015) found the pattern of accessibility of specific memories and specific future events approximated to the curve shown in Fig. 3. Fig. 3 is idealized data rather than actual data and in the study it is based on participants listed all the memories they could for each day going back five days and all the specific future events they could envisage plausibly.
occurring on each day for the five following days. The curve in Fig. 3 shows the (idealized) percentage of memories recalled/imagined for each day. It is in the window of the RIS that memories and imagined events are constructed and so become associated with the current representations of the recent past and near future.

One way to envisage the RIS is as a fish-eye lens in which items in the center of the lens are sharply defined and clear whereas items that lie progressively further away from the center become less focused and less distinct. In recent work in our laboratory (in progress) we have found that for events that progressively lie further in the future the more abstract, stylized, and schematic they become. They in effect lose their specificity and become more generic and scripted. Similarly, much of the more remote past becomes generic and inferred, with occasional specific memories. The RIS then is a window that extends the present moment some relatively short distance into the recent past and into the near future. In so doing we suggest that it is critical in supporting goal processing.

5. Memory accuracy: Correspondence & coherence

The concept of ‘accuracy’ in AM is complex. Memories might be thought of as falling in a 2-dimensional space of accuracy with ‘correspondence’ as one dimension and ‘coherence’ the other (see Conway, 2005). In this model each dimension would run from low to high creating four quadrants of: high correspondence-high coherence (unusually accurate memories), high correspondence-low coherence (memories of trauma might characterize this quadrant), low correspondence-high coherence (perhaps most memories fall somewhere in this quadrant), and low correspondence-low coherence (delusional and confabulated memories). The term correspondence refers to the case where a memory representation corresponds in some maximum way to a previously experienced event, in other words it is true to the event. Coherence refers to a memory representation that is coherent with other memory representations and self-beliefs (the conceptual self), in other words it is true to the self. Memories never fully correspond to our experience, only to parts of it, although they may be coherent with the self. Consider the following two memories:

When I was 14 I went to see the Beatles at the Buxton Pavilion Gardens. It was absolutely packed and me and my friend Jennifer were in the middle of the crowd. You couldn’t hear the music because there was so much screaming. I am only 5 feet tall and remember thinking I’m going to faint. Anyway I fainted and was carried to the stage. I came round just as I was being carried past the Beatles on stage. I was then thrown out of the back door, where there were many girls sitting crying or in hysterics. The door opened again and my friend was turfed out. As we stood there the Beatles started playing Twist and Shout and we ran back into the hall to see the end of the show.

A middle-aged man recalled his father distracting him when he was a young boy (about 4 years old) by asking who was the first man on the moon. He had been intensely interested in the moon landing when he was a young boy and this incident occurred while his father was on the telephone to his mother who had just given birth to his younger brother. He had a vivid and fond memory of his father placating him in this way, he was highly agitated by the birth, and in his memory he could ‘see’ his father on the telephone and almost ‘hear’ his voice. It was only decades later that he realized that his brother had been born in 1968, one year before the first moon landing.

The first memory illustrates a common feature of all memories: time compression. The memory contains some vivid details, some of which may correspond to the experience. Nonetheless, an event that took place over several hours is condensed into a short paragraph, in turn derived from a memory that took merely seconds to construct. Moreover, many details will have been added non-consciously and automatically, such as details of clothes worn, lighting, sound levels, the weather, etc. According to our model this would be a moderate correspondence-high coherence memory. The second vivid memory

---

2 Previously we had conceived correspondence-coherence as a single dimension. However, an anonymous reviewer of the present paper made the very helpful suggestion that we should change this to a 2-dimensional space and then all permutations of correspondence and coherence are possible. We thank the reviewer for their perceptive suggestion that we have now adopted.
may also contain some details that are true of the event, but the memory itself is of a supportive and caring father so it is a memory construction that is principally true of the self with which it is coherent. It is a low correspondence–high coherence memory.

In general, memories are *depictions* of the past constructed in the RIS from the complex underlying knowledge base. They fall somewhere in the correspondence-coherence space with all memories having aspects of both. One suggestion is that the mean of the distribution of memories in the correspondence-coherence 2D space is, for most people, closer to coherence than it is to correspondence. In this model it is possible for a memory to be simultaneously correct and incorrect, after all the owner of the 2nd memory above, did have a brother and a caring father, who most probably did phone the mother after the birth, etc. There are also other interesting properties of memories that emerge according to their degrees of correspondence and coherence. One of interest is the extent to which an individual’s memories although a mix of both correspondence and coherence are, overall, biased to one or other of the dimensions. Perhaps such biases contribute to determining different personality characteristics and types.

**6. Broken memories**

In concluding I will make a brief and highly selective overview of some of the ways in which memory can be impaired following brain damage, in psychological illness, and the effects of this on the SMS.

6.1. Confabulation

There is an extensive literature on confabulation resulting from various types of brain injury. I will consider two representative cases here. Patient O.P. was a 70-year-old woman who suffered lesions to the frontal and temporal lobes following an RTA (*Conway & Tacchi, 1996*). O.P. made a complete physical recovery but, on return home, her family complained of her constant ‘lies’. Her neuropsychological profile indicated spared IQ, language, and conceptual abilities. With some memory impairment and severe disturbance on frontal tests, O.P. could recall some memories very accurately but confabulated other memories about events relating to her family and these were nearly all focused on her grandson. They featured memories in which as a child he made implausible journeys to visit her, had major academic achievements, and later career achievements, all of which greatly pleased her. In fact he had been unsuccessful at school and in his career and lately had been unemployed for some years. He had little contact with her at any point in his life. Nonetheless, O.P.’s confabulations, or as Moscovitch (1995) termed them *honest lies*, were enduring and appeared to be highly emotionally positive for her. A case of coherence memories serving the purpose of providing a positive self-image during a time of considerable personal difficulty. Indeed, in many cases of confabulation following frontal damage it seems that the drive to make sense of the world persists but on the basis of false memories. For example, hospitalized frontal patients variably believe themselves to be in a hotel, at the office, waiting to play a round of golf, etc. Such meanings are not always positive and can be mixed with other confabulations that are emotionally negative, often highly negative (*Fotopoulou, Conway, Birchall, Griffiths, & Tyrer, 2007*). In either case, however, they reflect attempts to make sense of a world in which (false) memories constructed in the RIS drawn from knowledge in AM are constructed in ways that serve a disrupted working self. A working self in which personal beliefs and meanings are no longer grounded in and constrained by specific autobiographical memories.

Patient P.S. (*Hodges & McCarthy, 1993*) suffered a bilateral paramedian thalamic infarction resulting in a temporally graded retrograde amnesia (RA) and dense anterograde amnesia (AA). His RA stretched back to his early 20s and he believed himself to be a naval rating in the 2nd world war at home on shore leave – he was in fact a 60-year old businessman. Thus, the working self in trying to make meanings out of experience could only draw up on a RIS that had existed in some form when P.S. had been about 19 years of age. Current phenomena of daily life were interpreted through this RIS and, for example, PS was surprised by how much science fiction there was on the television and insisted in a nightly blackout. The case of P.S. illustrates very powerfully how the working self and indeed the RIS are constrained by the autobiographical knowledge they are able to access, that is by the cues control processes are able to use to create patterns of activation in the knowledge base, i.e. construct memories.

6.2. Psychological Illness

Baddeley and Wilson (1986) describe interviews with a series of schizophrenic patients. One striking feature of such patients is that their delusions were virtually always contradicted by other memories or autobiographical knowledge that, although wholly inconsistent with their beliefs no longer had the constraining power typical of a healthy autobiographical memory. In the 2D space of correspondence and coherence their memories would fall clearly in the quadrant denoting low correspondence and low coherence. For example, one patient, a man in his early 20’s, believed himself to be a famous rock guitarist, and confabulated memories of various ‘gigs’, while at the same time knew he had never learned to play the guitar and had never been to many of the places his delusions suggested. Another patient believed himself to be a Russian chess grandmaster, even though he knew he had never been to Russia and could recall being beaten at chess by other patients on his ward.
6.3. Posttraumatic stress disorder (PTSD)

Memories that derive from traumatic experiences often contain errors and distortion (Conway, Meares, & Standart, 2004). One patient, for example, had been sexually abused by his grandfather from ages of 4 to 12. He had a persistent and intrusive flashback in which he was naked in a bathroom being pushed against the radiator by his naked grandfather. In this intrusive fragment he saw himself from an observer perspective as he was now, a balding 35-year-old, and saw in the memory his grandfather as a frail 70-year-old. In fact it became clear during therapy that he had been 6 at the time and his grandfather was in his 40s. The false (coherence) memory served the function of obscuring the fact that he had been a helpless victim. Another PTSD patient who had witnessed the 9/11 attack on the World Trade Centre in New York had a vivid and deeply disturbing flashback in which he was flying in complete silence above the plane as it approached the tower. During therapy he eventually retrieved the true memory of standing in the crowd watching the plane strike the tower and then experienced the intense feelings of apprehension and fear that he had had at the time and also the later guilt that he had about the experience. Thus, distorted and false coherence memories can have a function of creating meaning that allow the working self to avoid negative self images and intense negative emotions.

7. Conclusion: The Modern View of Human Memory and Memory Accuracy

In the modern view of human memory memories are mental constructions. It is important to note that they are not reconstructions. They are not like videos, photographs or other recording media, even though they frequently contain mental imagery. They are transient constructions and although they may to some degree accurately represent the past they are time-compressed and contain many details that are inferred, consciously and non-consciously, at the time of their construction. Thus, all memories are to some degree false in the sense that they do not represent past experience literally. They can, of course, be wholly false but nonetheless be experienced as memories by a rememberer who may be unaware that the source of a memory is not experience but imagination. One of the main functions of memories is to generate meanings, personal meanings, that allow us to make sense of the world and operate on it adaptively. Memories are, perhaps, most important in supporting a wide range social interactions where coherence is predominant and correspondence often less central.

A strong implication of this view is that false memories far from being damaging to the individual can often be of considerable benefit, particularly in maintaining a coherent, confident, and positive self. Our time-compressed memories with their inferred and remembered details, their mixture of coherence and correspondence, serve this function perhaps for all of us. In closing, then, we might reflect on a hypothetical, if fairly extreme, example of this. Consider a person who has had an adverse childhood. Their confidence and sense of self-worth have been undermined by an emotionally, psychologically, and perhaps physically abusive parent. Their subsequent life has been a series of disasters marked by poverty, failed relationships, substance abuse, and disturbed children of their own. Through whatever means they gradually develop a set of memories of having been sexually abused by the parent. These memories are based in part on memories of actual occurrences of abuse that they experienced, although not on memories of sexual abuse. Details of the latter gradually become incorporated into some of their childhood memories by a process of imagination (see Garry, Manning, Loftus, & Sherman, 1996). They then come to experience these high coherence-low correspondence mental representations as memories and memories that in turn support a powerful personal meaning. Namely that the individual has been victimized (which in this hypothetical example indeed they have, but not sexually) and it is the effects of the victimization that have led to their disastrous life. Thus, the ‘memories’ provide the basis of an explanation for subsequent important parts of their life. This, of course, is merely an illustration (although perhaps not such an uncommon one, see Conway, 2013) of how memories (and recall that all memories are to some degree false) maybe of value in supporting the generation of explanatory beliefs about one’s life. We suggest it is something we all do.

References


