

# Evolution of working memory

Peter Carruthers<sup>1</sup>

Department of Philosophy, University of Maryland, College Park, MD 20742

Edited by Francisco J. Ayala, University of California, Irvine, CA, and approved March 20, 2013 (received for review January 21, 2013)

**Working memory (WM) is fundamental to many aspects of human life, including learning, speech and text comprehension, prospection and future planning, and explicit “system 2” forms of reasoning, as well as overlapping heavily with fluid general intelligence. WM has been intensively studied for many decades, and there is a growing consensus about its nature, its components, and its signature limits. Remarkably, given its central importance in human life, there has been very little comparative investigation of WM abilities across species. Consequently, much remains unknown about the evolution of this important human capacity. Some questions can be tentatively answered from the existing comparative literature. Even studies that were not intended to do so can nonetheless shed light on the WM capacities of nonhuman animals. However, many questions remain.**

attention | default network | episodic memory | mental rehearsal | primate

The nature of human working memory (WM) has been extensively investigated, with thousands of articles and books on the topic produced over the last half-century. Some of the main findings of this research will be outlined shortly. However, we know hardly anything about how WM evolved. For that (if we are to go beyond plausible speculation), we need detailed comparative studies. However, remarkably few such studies have been conducted, as we will see. Nevertheless, the emerging consensus about the nature of human WM allows us to frame a series of questions or alternative hypotheses concerning the possible differences between human and animal WM. Some of these can be answered, at least tentatively, from the results of existing work. However, they should also be used to frame and guide future comparative experiments.

## Working Memory in Humans

WM is the domain-general subsystem of the mind that enables one to activate and sustain (sometimes via active rehearsal) a set of mental representations for further manipulation and processing. The contents of working memory are generally thought to be conscious. Indeed, many identify the two constructs, maintaining that representations become conscious by gaining entry into WM (1). WM is generally thought to consist of an executive component that is distributed in areas of the frontal lobes working together with sensory cortical regions in any of the various sense modalities, which interact through attentional processes (2). It is also widely accepted that WM is quite limited in span, restricted to three or four chunks of information at any one time (3). Moreover, there are significant and stable individual differences in WM abilities between people, and these have been found to predict comparative performance in many other cognitive domains (4). Indeed, they account for most (if not all) of the variance in fluid general intelligence, or *g* (5).

The primary mechanism of WM is thought to be executively controlled attention (2, 6). It is by targeting attention at representations in sensory areas that the latter gain entry into WM, and in the same manner they can be maintained there through sustained attention. Attention itself is thought to do its work by boosting the activity of targeted groups of neurons beyond a threshold at which the information they carry becomes “globally broadcast” to a wide range of conceptual and affective systems throughout the brain while also suppressing the activity of

competing populations of neurons (1, 7, 8). These consumer systems for WM representations can produce effects that in turn are added to the contents of WM or that influence executive processes and the direction of attention. It is through such interactions that WM can support extended sequences of processing of a domain-general sort.

It is also widely accepted that WM and long-term (especially episodic) memory are intimately related. Indeed, many claim that representations held in WM are activated long-term memories (9). This might appear inconsistent with the claim that WM representations are attended sensory ones. However, the two views in part can be reconciled by noting that most models maintain that long-term memories are not stored in a separate region of the brain [although the hippocampus does play a special role in binding together targeted representations in other regions (10)]. Rather, information is stored where it is produced (often in sensory areas of cortex). Moreover, although attention directed at midlevel sensory areas of the brain appears to be necessary (and perhaps sufficient) for representations to enter WM, information of a more abstract conceptual sort can be bound into those representations in the process of global broadcasting (11). As a result, what figures in WM are often compound sensory-conceptual representations, such as the sound of a word together with its meaning or the sight of a face experienced as the face of one’s mother.

A final factor to stress is that WM is also intimately related to motor processes, probably exapting mechanisms for forward modeling of action that evolved initially for online motor control (12, 13). Whenever motor instructions are produced, an efferent copy of those instructions is sent to a set of emulator systems to construct so-called “forward models” of the action that should result. These models are built using multiple sensory codes (primarily proprioceptive, auditory, and visual), so that they can be aligned with afferent sensory representations produced by the action itself as it unfolds. The two sets of representations are compared, issuing in altered motor instructions if the action is failing to proceed as expected. These same systems are then used in the mental rehearsal of action, but with instructions to the muscles suppressed. The resulting sensory forward models, when targeted by attention, can gain entry into WM. Hence one can imagine oneself saying something and “hear” the result in so-called “inner speech,” or one can imagine oneself doing something and “see” or “feel” the results in visual or proprioceptive imagination.

Before we proceed to consider the evidence of WM in animals, it is important to distinguish WM from two other forms of memory with which it is sometimes conflated. One is sensory short-term memory, which can retain information in sensory cortices for around 2 s in the absence of attention. These representations

---

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “In the Light of Evolution VII: The Human Mental Machinery,” held January 10–12, 2013, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS Web site at [www.nasonline.org/evolution\\_vii](http://www.nasonline.org/evolution_vii).

Author contributions: P.C. performed research and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>E-mail: [pcarruth@umd.edu](mailto:pcarruth@umd.edu).

can give rise to priming effects without ever being conscious (14). (However, they can become conscious if attention is directed toward them before they expire. Consider the famous example of only noting the clock strike at the third chime while at the same time recalling the previous two strokes.) These sensory short-term memory representations can also be used for online guidance of action in the absence of attention (15). The contents of WM, in contrast, are attention-dependent and conscious and can be held in an active state for as long as attention is directed at them. (Note, however, that attention is quite sensitive to interference, so sustaining a representation in WM for an extended period is by no means easy.)

Some experimental results with animals that might be thought to support the existence of WM capacities are in fact best interpreted as tests of sensory short-term memory. Thus, consider the finding that chimpanzees and baboons can reliably recall a random sequence of spatial positions up to a limit of five to six items (or in the case of one animal, nine items) (16, 17). The temporal delays in these experiments are of the order of fractions of a second, with the animals' responses to the entire sequence generally being executed very swiftly over a period of around 2 s. So although these tasks might involve WM, the data can be accounted for in terms of sensory short-term memory alone.

The other contrast is with what is sometimes called in the human literature "long-term working memory" (18). Long-term working memory representations are those that are no longer among the active contents of WM (having fallen out of the focus of attention for too long), but which remain readily accessible to WM processes. Sometimes these representations have been recently activated from long-term memory, but sometimes they concern stimuli that were previously encoded into WM but were forgotten within a period of minutes. Long-term WM is thought to be important in speech and text comprehension, as well as underlying such phenomena as a bus-conductor's ability to know which of dozens of passengers on a bus have already paid for a ticket and which are newly arrived.

In this context it is important to note that numerous comparative studies of animals, such as those that use the radial-arm maze with rodents, use the term "working memory," when it is really a form of long-term WM that is being measured. The timescales involved, as well as the number of items that can be recalled, far exceed human WM abilities. Indeed, some writers are quite explicit that "working memory" in such studies should be defined as a memory that is used within a testing session (often lasting for minutes or hours) but not typically between testing sessions (such as the next day) (19, 20).

Empirically, WM can be distinguished from all forms of long-term memory by its sensitivity to attentional interference. Information sustained in WM will be lost if subjects are distracted and turn their attention fully to other matters. Long-term memories, in contrast, will merely decay at the normal rate in such circumstances. The authors of the study of serial-position memory in chimpanzees described above (16), for example, note that on some occasions the test subject was interrupted for a few seconds by a loud disturbance in a neighboring cage, but was nevertheless able to complete the sequence. Although the authors suggest that this behavior manifests the operation of WM, in fact it is unlikely (16). Undiminished performance following sustained and full distraction is a signature that long-term WM is involved.

### Working Memory in Animals

As we have seen, there are a number of aspects or components of normal WM function in humans, including capacities to sustain, rehearse, and manipulate active representations, with a signature limit of three to four items or chunks of information. We also know that WM is attention-dependent and hinges critically on capacities to resist interference from competing representations. Moreover, we know that WM plays a central role in many aspects

of intelligent human life. As a result, there are a range of possible positions that one can take concerning the comparative psychology of WM. These are listed below, organized roughly in terms of how great a gulf they envisage between the WM abilities of animals and ourselves. Thereafter they will be discussed in turn and evaluated in light of the available evidence.

- 1) Animals lack WM abilities altogether. They (like humans) have forms of sensory short-term memory that can retain reverberating information within sensory cortices for about 2 s following the removal of a stimulus, but they have no capacity to further sustain or refresh those representations.
- 2) Animals do have the capacity to sustain a representation of an object or event beyond the 2-s window of sensory short-term memory, but it is a very limited capacity—perhaps being restricted to one or two chunks in comparison with the three to four limit of humans.
- 3) Animals, like humans, can sustain three to four chunks of information in WM, but only in the absence of interference. Their abilities collapse (or are much weaker) when required to undertake a dual task or ignore intervening distractor items.
- 4) Animals have capacities to sustain representations that have been activated bottom-up, but they lack the capacity to activate a representation *ab initio*, using top-down attention to insert it into the global workspace. Basically, they lack imagination.
- 5) Animals can create and sustain representations in WM, but they lack any capacity to use mental rehearsals of action to generate contents for WM. [Some researchers use the term "rehearsal" to refer to the refreshing process that sustains short-term sensory representations in WM (21). I shall use it (as is commonly done) to refer to off-line rehearsals of action schemata that can be used to populate and sustain some of the contents of WM.]
- 6) Animals can create, sustain, and rehearse representations in WM, but they have limited capacities to manipulate those representations, transforming them and organizing them into effective problem-solving sequences in a controlled manner.
- 7) Animals have capacities to sustain, rehearse, and manipulate representations in WM much like our own. However, humans are unique in the extent to which they use their WM abilities. Specifically, humans frequently use WM in ways that are irrelevant to any current task (constituting the so-called "default network"), whereas animals' use of WM is always or generally task-oriented.
- 8) Animals have WM abilities much like our own and may even make chronic use of them. However, they differ in the sorts of representations that they can use in WM (in particular, lacking linguistic abilities, animals cannot generate inner speech), and their more limited conceptual repertoire limits the extent to which their WM performance can benefit from chunking.

We presently lack the evidence necessary for a thorough evaluation of any of these hypotheses beyond #1 and #8 of the list. However, there are data that bear directly on some of them, and some are more plausible than others on theoretical grounds. A sustained research effort by comparative psychologists is necessary for us to resolve these questions.

**1. No Capacity to Refresh and Sustain?** The most extreme position is to deny that animals have WM capacities at all. Animals nevertheless have forms of long-term memory as well as sensory short-term memory. But they have no capacity to refresh and sustain sensory activity in the absence of a stimulus or to keep representations active and available for longer time periods.

There are extensive data sufficient to exclude this possibility, much of it using match-to-sample or non-match-to-sample tasks. (Recall that data from animal experiments using the radial maze involve timescales too great to serve as direct tests of WM ability.) These tasks require an animal to remember the identity or location of a stimulus for more than a few seconds. By themselves these results of course cannot distinguish between the contributions of WM and long-term WM, and no doubt over extended intervals it will be long-term memory that is implicated. However, we also know from such studies that there are content-specific neurons in the prefrontal cortex that show sustained activity during retention intervals that are at least a few seconds long (22). Moreover, a great deal of what we know about the neurophysiology of human attentional and WM systems derives initially from work of this sort conducted with monkeys (23–25). So we can be confident that the mechanisms underlying WM performance in match-to-sample tasks are conserved across primates, and perhaps more widely.

In addition, numerous other studies have required animals to keep a representation of a target stimulus active beyond the 2-s window of sensory short-term memory. Some have used parallel object-displacement tests with apes and human children, with very similar results across all groups (26). Others have tested both apes and dogs to see whether they will continue to search for an item that they had seen placed in a “magic cup” after they had unexpectedly retrieved an item of a different sort with positive results (27).

The suggestion that basic WM capacities are quite widespread among animals receives additional support from neurobiology, given the tight connection between the WM system and episodic memory. (This will be discussed again in *Lack of Imagination?* below, where we review behavioral evidence of episodic-like memory in animals. Note here, however, that WM is the workspace within which episodic memories are activated and sustained by top-down attentional systems. And we have already noted that attentional networks are homologous among primates at least.) This is because the brain mechanisms subserving episodic-like memory are highly conserved among mammals. In particular, all mammals share homologous hippocampal and parahippocampal structures organized into homologous subregions, which have strong reciprocal connections to areas of the frontal cortex (28). These structures serve to integrate and store information about what occurred, where it occurred, and when it occurred (29). Indeed, even birds appear to share a similar, and at least partly homologous, network (28).

**2. One- or Two-Item Limit?** Some claim that nonhuman apes have a WM limit of two items, in contrast with the human WM limit of three to four chunks (30). However, this claim is based on a questionable analysis of the WM requirements of various tasks that apes cannot solve and assumes that failure does not result from other sources, such as a lack of understanding of physical forces and their effects. In contrast, experimental work with animals suggests that their WM limits may fall within the human range. Consider, for example, a test of serial recall of position conducted with a macaque monkey, modeled on tests that have been used with humans (31). The retention interval required in this test was about 4 s for the first item in the sequence, increasing to 11 s for the fourth, which places it squarely in the domain of WM. The monkey was successful in recalling the first three items in a sequence, but was at chance with the fourth. The experiment also demonstrated a very similar profile of recency, latency, and other effects commonly found with humans, suggesting that both species use a homologous WM mechanism with similar limits.

It should be stressed, however, that the work on human WM demonstrating that it has a capacity limit of three to four chunks [rather than Miller’s famous  $7 \pm 2$  (32)] has focused on the pure

memory-sustaining function of WM. Great care has been taken to exclude other strategies for maintaining representations in WM, such as covert mental rehearsal and informational chunking, which can extend its overall capacity still further (3). In the serial recall test just described, in contrast, the monkey may have used mental rehearsals of its planned movements to support its WM of the sequence of positions, thereby extending its pure memory-sustaining limits. This would be consistent with a claimed WM limit of one to two items.

Other data with animals suggesting WM limits in the human range are not so easily critiqued, however. For example, using paradigms that have previously been used with human infants, it has been shown that monkeys can track three to four items of food placed sequentially into one of two opaque containers (within which those items remain out of sight for a period of at least a few seconds). The monkeys reliably distinguish between containers that hold two versus three items, and also three versus four items, but not three versus five items (33). One might wonder why these data do not demonstrate that monkeys have a WM limit of seven (three items in one container and four in another) rather than four. The answer is that comparisons between containers benefit from chunking and do not just reflect raw retention limits. (A similar point holds for the infancy data.)

Similar tests have been conducted with horses, showing that they can distinguish between a bucket into which two apples have been placed and one containing three apples and fail to distinguish between buckets containing four apples and six apples, respectively (34). In such experiments, it seems unlikely that the animals could benefit from chunking because all of the items are of the same type. And it is likewise unclear how nonverbal forms of behavioral rehearsal could assist with the task (especially in the case of horses, whose repertoire of actions differs so widely from that of the human demonstrator). So the limit of three to four items revealed here seems most likely to reflect their pure WM retention capacity. However, until comparative psychologists use direct tests of simple WM retention abilities that can be conducted in parallel with adult humans, children, and members of various other species of animals, we will not be able to know for sure.

These results give rise to a puzzle, however. For, as noted earlier, variations in WM ability in humans are reliable predictors of fluid  $g$ . However, it seems that even monkeys have a WM span in the human range.\* This might lead one to expect similar general-learning abilities across all primates, which is manifestly false. A potential solution to the puzzle emerges when we note that the simple retention component of WM is *not* a reliable predictor of fluid  $g$  in humans (nor is it stable within a single individual across separate occasions of testing). Rather, only complex span tasks and so-called “n-back” tasks lead to stable results over time and are reliable predictors of  $g$  (37). (In a complex span test, one has to undertake some other task, such

\*A similar puzzle arises in the context of human development as it has been shown that WM capacity increases through the childhood years (35). In particular, 6- to 9-y-olds have a span of only two items or less in these experiments, whereas young adults have a span of three items. However, in other experiments, infants as young as 11 mo seem to already have an adult-like span of three items (36). One possible explanation is that speed of presentation differs between the two paradigms. In the experiments with children, the items-to-be-remembered are presented at a rate of one per second. In the experiments with infants, in contrast, presentation of each item takes a few seconds as the experimenter draws the infant’s attention to it, saying “Look at this.” Another possible explanation is that the infants participated in only a single trial, whereas the children had to keep attention to task across multiple presentations. Perhaps what changes through the childhood years is the capacity to maintain focused attention, rather than WM capacity as such. However, it may be that both of these explanations really amount to the same thing because the first explanation can be described in terms of the difference between directing attention toward an event (in accordance with task requirements) and having one’s attention drawn to an event.

as judging whether a simultaneously presented sentence makes sense or performing some simple mental arithmetic while also retaining an unrelated list in WM. In an *n*-back task, one has to keep track of the *n*-th prior item in a continually presented series, which requires one to resist interference from similar memories.) Moreover, at present it appears that it is training in *n*-back tasks—and not in simple span tasks—that issues long-term improvements in fluid *g* (38, 39, but see also ref. 40).

One possible construal of this set of findings is that there are no stable differences in simple span between people or across primate species. (As a result, simple span tests only measure noise contributed by endogenous factors or the environment.) All of the stable differences between people (and among species) may lie in the flexibility with which attention is allocated and the retention strategies used, as well as in the capacity to ignore sources of interference with targeted WM representations.

**3. Inability to Resist Interference?** There have been no controlled experiments comparing the abilities of humans and other animals to resist interference with WM representations. Clearly, the kinds of complex span tasks that have been used with humans are unsuitable for this purpose because most require linguistic abilities. However, there have been tests used with mice that tap into something quite similar. Some of these could be adapted for purposes of cross-species comparison.

Recent studies with mice have identified a general intelligence factor that explains about 40% of variance across a range of dissimilar learning tasks (41). Moreover, although this *g* factor is not significantly correlated with measures of simple WM retention, it is strongly correlated with performance in a more complex WM task, in which the animals have to resist interference from competing memories (42). In both cases the animals were first trained on two visually distinct radial-arm mazes located in the same room. In the test of WM retention, the animals were confined to the central compartment of one of the mazes for a fixed interval of 60 or 90 s, having made their first four correct choices before being allowed to complete their search. In the test of WM interference, in contrast, the animals were removed from the first maze, having made three correct choices and placed in the second maze; after three correct choices there, they were returned to the first maze until they had made another three correct choices, and so on. The fact that performance on the interference WM test but not on the retention WM test correlates with a measure of *g* in mice is suggestive of WM mechanisms homologous with those of humans.

One might question whether this and other experiments conducted in the same laboratory are genuinely measuring active WM rather than long-term WM. For how are we to know that the mice kept a representation of the arms already visited active in the focus of attention? Indeed, in experiments with rats using the eight-arm radial maze, rats typically show a near-perfect performance on the final four arms of the maze following delays of a number of hours after visiting the first four arms, enabling us to be quite confident that long-term memory is involved (20). On reflection, however, we can be sure that active WM is also used. So although the tests might not be suitable for measuring WM span (because both short-term and long-term WM are involved), they can enable us to draw conclusions about the relationship between WM and *g*.

Why should tests using interrupted search in a radial-arm maze involve interactions between short-term and long-term WM? When commencing search following an interruption, the animal will need to access long-term representations of the four arms previously visited, holding those in active WM long enough to select a fifth. And thereafter, for the final three choices, the animal will need to use spatial retrieval cues to access a long-term memory of each of the arms initially visited while keeping active in WM the immediately previous selections and while

orienting itself appropriately to make another choice. In addition, in the interference condition of the experiments described earlier (in which the mice are switched back and forth between two mazes), irrelevant memories will need to be suppressed, requiring the mice to pay careful online attention to the cues that individuate the arms of the two mazes. At the very least we can be confident that this task will place significant demands on the animals' use of selective attention, which is at the core of human WM abilities.

A subsequent study of correlations between WM abilities and *g* in mice attempted to determine the components of WM still further (43). It involved tests of WM retention time, WM retention capacity, as well as capacities for selective attention. The first experiment measured the temporal limits of the animals' capacity to recall which of the two arms in a T maze they had previously visited. The test of WM capacity used a nonspatial version of the radial-arm maze, in which cues attached to baited cups at the end of each arm were randomly shuffled following each choice, in such a way that the mice would need to keep in mind the cues (and which ones they had already selected) without relying on spatial position. Finally, the test of selective attention used two distinct discrimination tasks (one involving shapes and the other involving odors) that had initially been learned in separate contexts. During the test, the animals were presented with all cues of both kinds in one or the other of the two contexts, so that they would need to ignore one set of cues on which they had previously been trained in favor of the other. The results of this experiment were that retention time did not correlate with *g* at all and that WM capacity correlated moderately with *g*, whereas selective attention was strongly correlated with *g*. This, too, is what one might have predicted from what we know about human WM.

Perhaps the most impressive set of results from this series of studies with mice is the finding that WM training improves *g*, just as it appears to do in humans (44, 38). In the first of these experiments, animals who received training using two alternating radial-arm mazes scored significantly higher than controls on subsequent tests of general learning abilities and also scored higher on a test of selective attention. The second experiment then showed that it is the attentional component of WM training specifically that leads to an improvement in *g*. This experiment used three groups of mice. One group received training in two alternating and visually similar radial-arm mazes located within the same room, which would require the mice to attend to minor differences in cues provided by spatial context to discriminate the arms of the two mazes. A second group also received training on two alternating radial-arm mazes, but this time located in separate rooms, thus placing fewer attentional demands on the animals. The third group was a control and received no WM training. The findings were that the attentionally demanding group showed the greatest increase in *g* and the second group also displayed significant improvement relative to controls.

Taken together, this series of findings with mice suggests that WM abilities in this species are heavily dependent on attentional capacities (just as they are in humans) and that mice not only have a simple capacity to retain salient information beyond the temporal window of sensory short-term memory, but also (like humans) can do so in the face of interference. It may be, then, that the basic structure of WM is at least homologous across all mammals. However, we do not know to what extent (if at all) capacities to direct and control attention and to resist interference differ between humans and other mammals. Given that such capacities are aspects of executive function, and that humans are generally supposed to excel at executive function tasks, one might predict significant differences. However, the situation cries out for direct tests of attentional abilities and complex WM capacities across species.

**4. Lack of Imagination?** There are two basic ways in which offline representations can gain entry into WM. One is through mental rehearsals of action, which are discussed in section 5 below. The other is through top-down executive-attentional processes. One can search for, and activate into WM, a visual image of one's mother's face or an auditory image of the sound of her voice, for example. However, one can also search for and activate a specific episodic memory of one's graduation or one's most recent birthday dinner. It seems most likely that these two forms of ability are paired together. However, it would be possible to claim that a creature can have a capacity for generic semantic imagery without being capable of episodic memory, perhaps because representations of specific episodes are never stored in memory at all. So even if animals are incapable of mental time travel (including episodic remembering), as some have claimed (45), this would fail to show that they are incapable of using attentional resources to generate imagistic contents for WM in an offline manner. If animals are capable of episodic remembering, in contrast, they then will surely also be capable of generic imagery because it is hard to see what more might be required for the latter than is already present in the former.

Most tests of mental time travel in animals have focused on prospection of the future (discussed in section 5 below). However, there have also been experiments with corvids showing these birds to be at least capable of recalling and reasoning appropriately from the what, where, and when components of episodic memory (46). Admittedly, it does not follow that the birds are experientially projecting themselves back into specific episodes of food caching. However, it does at least seem likely that they are activating into WM episodic-like representations of types of food and their locations, together with some sort of representation of elapsed time. At any rate, this is how humans would solve a problem of this sort if compelled to do so nonverbally. This consideration would provide a stronger argument, of course, if corvids were not so evolutionarily distant from us. However, despite this distance, we noted earlier that birds possess brain networks that are similar to, and at least partly homologous with, those that support episodic memory in humans and other mammals (28). Moreover, experiments with rats show that they, too, form tightly integrated what, where, and when representations (47, 48). Such data suggest that episodic-like memory representations are widespread among animals. However, in any case it seems that the animals must at least be capable of activating representations into WM using top-down attentional control.

Recall, moreover, the experiments with rodents using interrupted search of a radial-arm maze, discussed in section 3 above. Although there is nothing in the data to suggest that in the second phase of the experiments the animals are accessing episodic memories of their earlier visits to some of the arms of the maze, they will surely at least be activating a semantic representation of some sort. For example, it might be a representation of an arm as being empty of any reward. In humans, such a memory would need to be searched for using a combination of environmental cues and top-down attentional control, resulting in that representation being activated into WM. It is therefore reasonable to assume that the same is true of rodents.

There are tentative grounds, then, for thinking that other animals are capable of top-down activation of representations for use in WM. Further grounds are discussed in section 5 below, because it is unlikely that the use of WM for prospection depends solely on activation of motor schemata without any enrichment from semantic or episodic memory. Indeed, we know that long-term memory systems and capacities for prospection are tightly linked, with the hippocampus being heavily implicated in each (49). In fact, some have argued that the structure of long-term memory systems has been specifically adapted and shaped in the service of prospective reasoning (50).

Moreover, one might think, on purely theoretical grounds, that any creature capable of top-down attentional selection of stimuli should also be capable of top-down activation of similar representations in an offline manner. For as we noted earlier, attention operates by boosting the neural activity of some groups of neurons while simultaneously suppressing the activity of competing populations, resulting in global broadcast of the information encoded in the former set. The same mechanisms should then be capable of operating in the presence of background levels of neural activation in the absence of an external stimulus, resulting in endogenous activation of representations in the global workspace.

**5. Inability to Mentally Rehearse Action?** Evidence of mental rehearsal of action comes from studies of long-term planning in animals. We know that in humans such planning is conducted in large part through rehearsal of alternative actions, with people responding affectively to the WM representations that result (51, 52). Although there is powerful evidence of future planning in corvids (53, 54), I shall focus on data from primates, where the argument for homologous underlying mechanisms is strongest.

One study has carefully documented the behavior of an alpha male chimpanzee in an open-plan zoo (55, 56). He began to collect and store piles of stones early in the morning to throw at zoo visitors later in the day as part of an aggressive threat display. When the zoo keepers responded by removing his stashes each day before zoo opening time to prevent this, he proved quite adept at concealing his stashes and at manufacturing projectiles afterward by breaking off pieces of brittle concrete from the walls in his enclosure. Note that at the times when he collected and concealed his stashes he was in a calm state, in the absence of the stimuli (human visitors) that would provoke his rage later. Such behavior in a human would likely be caused by imagining the later presence of the audience and mental rehearsal of the actions involved in grasping and throwing projectiles, issuing in a positive affective response that would in turn motivate the collection of some stones. It is reasonable to assume that similar processes took place in the mind of the chimpanzee.

Experimental data with chimpanzees point toward the same conclusion. In one experiment, chimpanzees not only selected and carried with them to their sleeping quarters a tool that they would need the next day to access a desired reward, but also remembered to bring it back with them on their return (57). In a conceptual replication of this experiment by another laboratory, chimpanzees again selected a tool needed to retrieve a later reward and remembered to bring the tool with them when returning (58). Moreover, the animals were able to resist a smaller current reward (a grape), choosing instead the tool that would get them a more valued reward later (a container of juice). In addition, when presented with a number of unfamiliar objects (while being prevented from handling them), they reliably selected and took with them the one best suited to obtain the future reward. Note that humans would solve a task of this sort by mentally rehearsing some actions directed toward the juice container involving the various objects, noting which ones could be successful.

This evidence from captive chimpanzees is fully consistent with what we know of the behavior of chimpanzees in the wild. For example, chimpanzees in the Congo regularly harvest termites from both above-ground and subterranean nests, each of which requires a distinct set of tools. The subterranean nests, in particular, require a sharp stout puncturing stick, which is always made from the branches of a particular species of tree. The chimpanzees never arrived at the site of a subterranean nest without bringing such a stick with them, unless one had previously been left at the site. And this was true even though the nearest appropriate tree was tens of meters away in the forest, from which point the nest site could not be seen (59). Such behavior in humans would involve imagination of the target together with mental rehearsal of

the actions needed to acquire it, which would both remind and motivate one to deviate from one's path to find an appropriate species of tree.

The behavioral data suggest, then, that other apes (at least) are capable of mentally rehearsing actions and that they do so for purposes of future planning, just as humans do. However, at present the argument for this conclusion is one of analogy, assuming that similar forms of behavior across closely related species should be explained in terms of similar underlying processes. Evidence of a more direct sort would be quite welcome. In particular, we need experimental paradigms that can be matched across species, whose parameters can be varied in parallel to see whether performance profiles respond similarly also. A positive outcome would provide much stronger evidence of homologous processes.

**6. Limited Manipulative Abilities?** In one sense, the manipulative component of WM consists of an ability to organize and control sequences of representations in a task-relevant manner. The evidence of future planning in apes and corvids suggests that they are capable of doing just that. In another sense, however, manipulation involves targeting an image with a mentally rehearsed action, thereby transforming it. This has been extensively studied in humans using the visual rotation paradigm (11). Participants are presented with two shapes of varying orientation and are asked to judge whether or not the shapes are the same. People solve these tasks by mentally rotating the image of one shape to match the orientation of the other and answering depending on whether or not the result is a fit. Among the classic findings in this literature are that participants take longer to judge shapes whose orientations are further apart from one another, suggesting that the movement of the initial image through the intervening space takes time.

What we know from brain-imaging and transcranial magnetic stimulation studies using the visual rotation paradigm is that activity in the motor or premotor cortex precedes and causes the subsequent transformation of the visual image (60). It seems that people imagine acting on the shape represented in one of the images, initiating offline an action of twisting it with one's hand, for example, thereby causing the represented shape to change through the process of forward modeling of the action. One might wonder, then, whether animals have similar capacities. Studies conducted with baboons and sea lions suggest that they do, with the animals showing larger differences in reaction time to images that would need to be rotated through larger arcs to secure a match, just as humans do (61, 62). However, to justify claiming that the processes are homologous it would be important to know whether motor-control areas of the animals' brains are likewise involved in the process.

Similar conclusions are supported by studies of problem solving and insight in apes. For example, confronted by a peanut at the bottom of a glass container that is too deep to reach into (and which is strapped to the bars of the cage), some animals will hit upon the strategy of collecting water in their mouths and spitting it into the container until the peanut floats to the top (63, 64). (The same task was presented to human children, with similar rates of success among 4- and 6-y-old children, but with more frequent achievement among 8-y-olds.) To arrive at the solution to this problem, one needs to mentally rehearse an action of putting water into the container, thereby transforming one's mental representation of the position of the peanut and enabling one to predict that iterated performance of the action will permit one to reach it successfully. However, once again the argument for homologous processes here is only one of analogy.

**7. Rarity of Use?** Even if the WM capacities of animals are comparable to those of humans in all major respects, it may be that animals make use of WM only when confronted with specific

practical, learning, or reasoning problems. Humans, in contrast, make frequent use of WM in ways that are irrelevant to any current task, thereby constituting the default network (65, 66).<sup>†</sup> Even when we are not confronted with a task our minds will be occupied with fantasies, episodic memories, imagined social situations, imagined conversations, snatches of song, and so on, all of which heavily involve WM. Indeed, even when humans are engaged in a task, they are apt to slip into so-called "mind wandering," in which WM is populated with representations unrelated to the task demands (68).

There is little comparative data bearing directly on this question. However, the suggestion that humans may be unique in this respect is at least consistent with the vastly greater extent of human creativity, innovation, and long-term planning. Much of the time that humans spend mind wandering is occupied with reviewing and exploring future scenarios and anticipating future problems or successes. Moreover, there is evidence that mind wandering is significantly correlated with creativity, involving, as it does, defocused attention combined with executive control and selection (69). It has also been suggested that the uniquely human disposition to engage in pretend play in childhood is an adaptation for increased creativity in adulthood, encouraging us to use WM for purposes of creative scenario building (70).

Data suggesting that mind wandering may not be uniquely human come from a study comparing default network activity in humans and chimpanzees (71). Similar regions of the brain displayed greater activity at rest in both species, including in the medial prefrontal cortex and posterior cingulate cortex, suggesting that chimpanzees, too, spend much of their time ruminating when not engaged in other tasks. These data need to be treated with caution, however, because default-mode networks overlapping those of humans have been found in both monkeys and rodents under conditions of general anesthesia (72, 73). Therefore default-mode activity does not entail conscious mind wandering of the sort that would implicate the resources of WM. Rather, the explanation for these findings may be that the main components of the default network (especially medial regions of both the prefrontal and the parietal cortex) are important connecting hubs in the neural architecture of the brain, serving to link together other more modular regions (74). As such the prefrontal and the parietal cortex will generally exhibit greater neural activity than the regions that they connect, just as airports that serve as major hubs show greater flight activity than others. In humans, we know that these default-network hubs play an important role in mind wandering. However, it does not follow that any animal with similar brain connectivity will also make use of its WM when at rest to replay the past and explore the future in the ways in which humans do.

It might be proposed that we have direct evidence of such replay activity in rats. When at rest, or during pauses in exploration of a track, place cells in the rat hippocampus fire in sequences corresponding to portions of the route already traveled or about to be traveled (75). However, although these firing sequences take place over intervals that are linearly related to the distances represented, firing rates are very fast in comparison with the rat's normal rate of motion (corresponding to rates of

<sup>†</sup>Brain-imaging studies of the default network rarely find activity in sensory cortices of the sort that one would expect to accompany WM use. In part, this may be an artifact of the subtraction methodology involved in these studies because the paired nondefault conditions will generally involve attention to some perceptually presented task. However, it may also be because different participants (or the same participant at different times) are using the resources of distinct sense modalities, engaging in inner speech on some occasions and visual imagery on others. What is generally agreed is that default-mode operation consists of episodic remembering, prospection of the future, and so on, which are known to make use of WM. And indeed, paired perception and imaging tasks in two distinct sense modalities (hearing and vision) show both a common core network implicated in each (which largely overlaps with the default-mode network) and modality-specific activity in midlevel sensory areas that varies by condition (67).

about 8 m/s). In fact, the rate of “mental travel” is 15–20 times faster than actual travel. This contrasts sharply with the finding that, when humans imagine walking across a room, their imagined journey takes place at approximately the same speed as an actual journey (76). This suggests that the processes are not homologous across the two species and may serve quite different functions. Indeed, it is generally thought that rapid place-cell firing probably plays a role in the consolidation of memory (and, as such, is likely to take place in humans as well as in rodents).

It seems, then, that at present there is no real evidence to counter the suggestion that humans are unique in making frequent use of WM for purposes of rumination and mind wandering. However, this suggestion is supported (albeit quite weakly) by a theoretical inference from differences in long-term planning and creativity.

**8. More Limited Behavioral and Conceptual Resources.** Even if animals have WM capacities that are in all respects like our own, and likewise make chronic use of them, we can be confident that they are systematically different from us in the contents that figure in their WM. The primary reason for this is that only humans are capable of speech. This means that there is an entire range of actions (namely, speech actions) that only humans can mentally rehearse. In addition, the vastly greater conceptual repertoire possessed by humans (in part resulting from previous speech communication) will mean that humans have available many more ways in which to chunk information in WM, thereby extending the latter’s scope and flexibility.

It is in these terms that we can characterize the unique character of so-called “system 2” reasoning and decision making in humans. Psychologists who study human reasoning have increasingly converged on the hypothesis that we use two distinct sets of processes when doing so (77–79). System 1 is swift, unconscious, and intuitive and is thought to be largely shared with other animals. System 2 is reflective, serial, and slow, and its operations are largely conscious, using the limited resources of WM. Many (but by no means all) system 2 processes use mental rehearsals of sentences and phrases in inner speech, so in this

respect system 2 is uniquely human. Moreover, given that WM and fluid *g* largely coincide, differences in WM capacities explain a significant portion of the variance between people in tests of their reasoning abilities, with the remainder of the variance being accounted for by differences in people’s disposition to stop and reflect before answering and in their knowledge of norms of reasoning, or their “mindware” (80).

If the animal studies reviewed above have been correctly interpreted, then system 2 as such will not be uniquely human. For any animal engaged in prospection, envisaging and responding affectively to the consequences of the various actions open to it (which are mentally rehearsed in sequence) will qualify as engaging in system 2 processing. What is unique to humans is our ability to vastly extend the topics and forms of reflective thinking in which we can engage by virtue of our capacity for mental rehearsal of speech.

## Conclusion

We can be confident that other primates, at any rate, have WM systems in many respects homologous with our own. We can be just as confident that humans are unique in some of the uses that they make of WM, specifically of inner speech. However, between these two items of knowledge there is a large space of possibilities about which little is known for sure. It seems likely, on current evidence, that other primates (and perhaps all mammals) have pure retention abilities whose limits are similar to those of humans. Moreover, whereas humans are by no means unique in having a capacity for prospection and future planning using WM, it seems likely that humans excel in their abilities to withstand interference and to deploy attention and rehearsal in flexible ways to maintain and manipulate representations in WM. In addition, there is some reason to suspect that humans may be unique in making frequent task-independent use of their WM abilities. However, until there is a sustained effort by comparative psychologists to devise and carry out matching tests of WM ability involving humans and various other species of animal, many of these claims must remain at least partly speculative.

- Baars BJ (2002) The conscious access hypothesis: Origins and recent evidence. *Trends Cogn Sci* 6(1):47–52.
- Postle BR (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139(1):23–38.
- Cowan N (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci* 24(1):87–114, discussion 114–185.
- Engle R (2010) Role of working-memory capacity in cognitive control. *Curr Anthropol* 51:517–526.
- Kane MJ, Hambrick DZ, Conway AR (2005) Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychol Bull* 131(1):66–71, author reply 72–75.
- Cowan N, et al. (2005) On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognit Psychol* 51(1):42–100.
- Gazzaley A, Cooney JW, McEvoy K, Knight RT, D’Esposito M (2005) Top-down enhancement and suppression of the magnitude and speed of neural activity. *J Cogn Neurosci* 17(3):507–517.
- Knudsen EI (2007) Fundamental components of attention. *Annu Rev Neurosci* 30: 57–78.
- Unsworth N, Engle RW (2007) The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychol Rev* 114(1):104–132.
- Squire LR (1992) Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychol Rev* 99(2):195–231.
- Kosslyn S (1994) *Image and Brain* (MIT Press, Cambridge, MA).
- Wolpert DM, Ghahramani Z (2000) Computational principles of movement neuroscience. *Nat Neurosci* 3(Suppl):1212–1217.
- Jeannerod M (2006) *Motor Cognition* (Oxford Univ Press, Oxford).
- Dehaene S, Changeux JP, Naccache L, Sackur J, Sergent C (2006) Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends Cogn Sci* 10(5):204–211.
- Milner D, Goodale M (1995) *The Visual Brain in Action* (Oxford Univ Press, Oxford).
- Inoue S, Matsuzawa T (2007) Working memory of numerals in chimpanzees. *Curr Biol* 17(23):R1004–R1005.
- Fagot J, De Lillo C (2011) A comparative study of working memory: Immediate serial spatial recall in baboons (*Papio papio*) and humans. *Neuropsychologia* 49(14): 3870–3880.
- Ericsson KA, Kintsch W (1995) Long-term working memory. *Psychol Rev* 102(2):211–245.
- Dudchenko PA (2004) An overview of the tasks used to test working memory in rodents. *Neurosci Biobehav Rev* 28(7):699–709.
- Shettleworth S (2010) *Cognition, Evolution, and Behavior* (Oxford Univ Press, Oxford), 2nd Ed.
- Jonides J, et al. (2008) The mind and brain of short-term memory. *Annu Rev Psychol* 59:193–224.
- Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14(3):477–485.
- Goldman-Rakic PS, Funahashi S, Bruce CJ (1990) Neocortical memory circuits. *Cold Spring Harb Symp Quant Biol* 55:1025–1038.
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol* 77(1):24–42.
- Baluch F, Itti L (2011) Mechanisms of top-down attention. *Trends Neurosci* 34(4): 210–224.
- Barth J, Call J (2006) Tracking the displacement of objects: A series of tasks with great apes and young children. *J Exp Psychol* 32:239–252.
- Bräuer J, Call J (2011) The magic cup: Great apes and domestic dogs (*Canis familiaris*) individuate objects according to their properties. *J Comp Psychol* 125(3):353–361.
- Allen T, Fortin N (2013) The evolution of episodic memory. *Proc Natl Acad Sci USA* 110:10379–10386.
- Eichenbaum H (2013) Memory on time. *Trends Cogn Sci* 17(2):81–88.
- Read DW (2008) Working memory: A cognitive limit to non-human primate recursive thinking prior to hominid evolution? *Evol Psychol* 6:676–714.
- Botvinick MM, et al. (2009) An analysis of immediate serial recall performance in a macaque. *Anim Cogn* 12(5):671–678.
- Miller GA (1956) The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychol Rev* 63(2):81–97.
- Hauser MD, Carey S, Hauser LB (2000) Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proc Biol Sci* 267(1445):829–833.
- Uller C, Lewis J (2009) Horses (*Equus caballus*) select the greater of two quantities in small numerical contrasts. *Anim Cogn* 12(5):733–738.
- Cowan N, AuBuchon AM, Gilchrist AL, Ricker TJ, Saults JS (2011) Age differences in visual working memory capacity: Not based on encoding limitations. *Dev Sci* 14(5): 1066–1074.

36. Feigenson L, Carey S (2005) On the limits of infants' quantification of small object arrays. *Cognition* 97(3):295–313.
37. Engle R (2010) Role of working-memory capacity in cognitive control. *Curr Anthropol* 51:S17–S26.
38. Jaeggi SM, Buschkuhl M, Jonides J, Perrig WJ (2008) Improving fluid intelligence with training on working memory. *Proc Natl Acad Sci USA* 105(19):6829–6833.
39. Jaeggi SM, Buschkuhl M, Jonides J, Shah P (2011) Short- and long-term benefits of cognitive training. *Proc Natl Acad Sci USA* 108(25):10081–10086.
40. Chooi W-T, Thompson L (2012) Working memory training does not improve intelligence in healthy young adults. *Intelligence* 40:531–542.
41. Matzel LD, et al. (2003) Individual differences in the expression of a "general" learning ability in mice. *J Neurosci* 23(16):6423–6433.
42. Kolata S, et al. (2005) Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiol Learn Mem* 84(3):241–246.
43. Kolata S, Light K, Grossman HC, Hale G, Matzel LD (2007) Selective attention is a primary determinant of the relationship between working memory and general learning ability in outbred mice. *Learn Mem* 14(1):22–28.
44. Light KR, et al. (2010) Working memory training promotes general cognitive abilities in genetically heterogeneous mice. *Curr Biol* 20(8):777–782.
45. Suddendorf T, Corballis MC (2007) The evolution of foresight: What is mental time travel, and is it unique to humans? *Behav Brain Sci* 30(3):299–313, discussion 313–351.
46. Clayton NS, Yu KS, Dickinson A (2003) Interacting Cache memories: Evidence for flexible memory use by Western Scrub-Jays (*Aphelocoma californica*). *J Exp Psychol Anim Behav Process* 29(1):14–22.
47. Ergorul C, Eichenbaum H (2004) The hippocampus and memory for "what," "where," and "when" *Learn Mem* 11(4):397–405.
48. Babb S, Crystal J (2005) Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learn Motiv* 36:177–189.
49. Buckner RL (2010) The role of the hippocampus in prediction and imagination. *Annu Rev Psychol* 61:27–48, C1–C8.
50. Schacter DL, Addis DR, Buckner RL (2007) Remembering the past to imagine the future: The prospective brain. *Nat Rev Neurosci* 8(9):657–661.
51. Damasio A (1994) *Descartes' Error* (G P Putnam's Sons, New York).
52. Gilbert DT, Wilson TD (2007) Propection: Experiencing the future. *Science* 317(5843):1351–1354.
53. Correia SP, Dickinson A, Clayton NS (2007) Western scrub-jays anticipate future needs independently of their current motivational state. *Curr Biol* 17(10):856–861.
54. Taylor AH, Elliffe D, Hunt GR, Gray RD (2010) Complex cognition and behavioural innovation in New Caledonian crows. *Proc Biol Sci* 277(1694):2637–2643.
55. Osvath M (2009) Spontaneous planning for future stone throwing by a male chimpanzee. *Curr Biol* 19(5):R190–R191.
56. Osvath M, Karvonen E (2012) Spontaneous innovation for future deception in a male chimpanzee. *PLoS ONE* 7(5):e36782.
57. Mulcahy NJ, Call J (2006) Apes save tools for future use. *Science* 312(5776):1038–1040.
58. Osvath M, Osvath H (2008) Chimpanzee and orangutan forethought: Self-control and pre-experience in the face of future tool use. *Anim Cogn* 11:661–674.
59. Sanz C, Morgan D, Gulick S (2004) New insights into chimpanzees, tools, and termites from the Congo Basin. *Am Nat* 164(5):567–581.
60. Ganis G, Keenan JP, Kosslyn SM, Pascual-Leone A (2000) Transcranial magnetic stimulation of primary motor cortex affects mental rotation. *Cereb Cortex* 10(2):175–180.
61. Vauclair J, Fagot J, Hopkins W (1993) Rotation of mental images in baboons when the visual input is directed to the left hemisphere. *Psych Sci* 4(2):99–103.
62. Mauck B, Dehnhardt G (1997) Mental rotation in a California sea lion (*Zalophus californianus*). *J Exp Biol* 200(Pt 9):1309–1316.
63. Mendes N, Rakoczy H, Call J (2008) Ape metaphysics: Object individuation without language. *Cognition* 106(2):730–749.
64. Hanus D, Mendes N, Tennie C, Call J (2011) Comparing the performance of apes and human children in the floating peanut task. *PLoS ONE* 6:e19555.
65. Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: Anatomy, function, and relevance to disease. *Ann N Y Acad Sci* 1124:1–38.
66. Spreng RN, Mar RA, Kim AS (2009) The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *J Cogn Neurosci* 21(3):489–510.
67. Daselaar SM, Porat Y, Huijbers W, Pennartz CM (2010) Modality-specific and modality-independent components of the human imagery system. *Neuroimage* 52(2):677–685.
68. Mason MF, et al. (2007) Wandering minds: The default network and stimulus-independent thought. *Science* 315(5810):393–395.
69. Baird B, et al. (2012) Inspired by distraction: Mind wandering facilitates creative incubation. *Psychol Sci* 23(10):1117–1122.
70. Picciuto E, Carruthers P (2013) The origins of creativity. *The Philosophy of Creativity*, eds Paul E, Kaufman S (Oxford Univ Press, Oxford).
71. Rilling JK, et al. (2007) A comparison of resting-state brain activity in humans and chimpanzees. *Proc Natl Acad Sci USA* 104(43):17146–17151.
72. Vincent JL, et al. (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447(7140):83–86.
73. Lu H, et al. (2012) Rat brains also have a default mode network. *Proc Natl Acad Sci USA* 109(10):3979–3984.
74. Sporns O (2011) *Networks of the Brain* (MIT Press, Cambridge, MA).
75. Davidson TJ, Kloosterman F, Wilson MA (2009) Hippocampal replay of extended experience. *Neuron* 63(4):497–507.
76. Decety J, Jeannerod M, Prablanc C (1989) The timing of mentally represented actions. *Behav Brain Res* 34(1–2):35–42.
77. Evans J, Frankish K, eds (2009) *In Two Minds: Dual Processes and Beyond* (Oxford Univ Press, Oxford).
78. Kahneman D (2011) *Thinking, Fast and Slow* (Farrar, Strauss and Giroux, New York).
79. Evans J (2010) *Thinking Twice* (Oxford Univ Press, Oxford).
80. Stanovich K (2009) *What Intelligence Tests Miss* (Yale Univ Press, New Haven, CT).