The neuroscience of working memory capacity and training

Christos Constantinidis[1] and Torkel Klingberg[2]

Abstract | Working memory — the ability to maintain and manipulate information over a period of seconds — is a core component of higher cognitive functions. The storage capacity of working memory is limited but can be expanded by training, and evidence of the neural mechanisms underlying this effect is accumulating. Human imaging studies and neurophysiological recordings in non-human primates, together with computational modelling studies, reveal that training increases the activity of prefrontal neurons and the strength of connectivity in the prefrontal cortex and between the prefrontal and parietal cortex. Dopaminergic transmission could have a facilitatory role. These changes more generally inform us of the plasticity of higher cognitive functions.

Working memory (WM) was first proposed to be the memory for plans of future action[1]. The field of cognitive psychology has underlined the importance of the ability to manipulate the information kept in WM and the role of WM for complex mental abilities such as learning and reasoning[2]. Lesion studies first associated WM with the prefrontal cortex (PFC)[3,4], and this link was later confirmed by electrophysiological recordings that revealed neural correlates of WM in the PFC of monkeys[5,6]. In an early summary of the then-available neurophysiological and cognitive research, WM was described as the "capacity to guide behaviour by representation of stimuli rather than by the stimuli themselves", thus introducing "the possibility that concepts and plans can govern behaviour" (REF. 7).

One of the central limitations of human cognition is the restricted amount of information that can be kept in WM[8]. Individual differences in WM capacity are associated with variation in several important abilities, including control of attention[9], non-verbal reasoning ability[10] and academic performance[11]. A lower WM capacity is also a feature in many clinical populations, including individuals with schizophrenia, stroke, traumatic brain injury and attention deficit-hyperactivity disorder (ADHD). Finding the neural mechanisms of this limitation is therefore a key goal in cognitive neuroscience.

It was traditionally assumed that WM capacity is an immutable individual characteristic, but research at the beginning of the 2000s showed that the WM capacity of children and young adults could be increased by using computerized training that allows more extensive training (often more than 12 hours)[12–14]. After such training, performance increases not only for the trained tasks but also for WM tasks that were not part of the training; that is, there is a transfer of performance improvement from one task, or set of tasks, to another. The amount of improvement transferred to non-trained tasks, relative to control groups, is estimated by three recent meta-analyses to be around 0.6 standard deviations[15–17]. Ongoing research aims to elucidate the extent to which the improvements also translate to various other abilities, such as academic abilities, that statistically correlate with WM capacity. Indeed, transfer to reasoning ability has been reported to be small, with inconsistent findings. More-encouraging transfer results are seen for attention: several randomized, controlled trials, some of them including children with ADHD, show decreases in inattentive behaviour in everyday life after WM training[13,18–21], consistent with the close functional relationship between neural systems for WM and attention[17]. Importantly, the transfer from trained to non-trained WM tasks shows that WM, a key cognitive ability, is malleable. Studying the neural basis of this change could provide important information on brain plasticity relating to cognitive functions.

The neural correlates of WM training have been investigated in studies involving functional MRI (fMRI), electroencephalography (EEG), transcranial magnetic stimulation, positron-emission tomography (PET) and genetic analysis. Electrophysiological studies in animal models show the effects of training at the neuronal, synaptic and receptor levels, and combining this knowledge could lead to an understanding of the mechanisms underlying differences in WM capacity and plasticity.

In this Review, we examine the neural bases that underlie WM processes in non-human primates and humans, and how these processes have been incorporated into computational neural network models. These fields encompass different tasks, methods and definitions, but they also have

[1]*Department of Neurobiology and Anatomy, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157, USA.*
[2]*Deptartment Neuroscience, Karolinska Institutet, 171 77 Stockholm, Sweden.*

Correspondence to T.K.
torkel.klingberg@ki.se

doi:10.1038/nrn.2016.43
Published online 26 May 2016

Box 1 | **Studies of working memory in rodents**

In addition to neurophysiological studies in non-human primates, much research on the neural bases of working memory (WM) has been carried out in rodents. One of the tasks that is suggested to be a rodent equivalent to human and macaque WM tasks is the eight-arm radial maze, in which information about the location visited needs to be remembered across consecutive choices. Lesions of the prelimbic region, a part of the cortex that is suggested to be a homologue of the primate prefrontal cortex (PFC), impair performance on WM tasks and attention-demanding tasks[123]. Persistent activity during the delay in WM tasks has been demonstrated in rodents, at least during WM intervals that require preparation for a motor response[124]. An alternative mechanism for the maintenance of WM in rodents involves transient activation of populations of neurons that come 'online' at different time points in order to span the entirety of a delay period[125].

WM training has been shown to improve performance across a range of tasks in rodents[126]. Such training induces long-lasting molecular changes, including chromatin remodelling and expression of immediate early genes, such as the gene encoding activity-regulated cytoskeleton-associated protein (ARC)[127,128].

The effects of training on WM-related activity have only recently been addressed in rodents. Increased prelimbic activity has generally been reported following various types of task learning that require WM[129]. Training also results in enhanced dopamine D1 receptor-mediated neuronal activation in the prelimbic cortex[114].

commonalities. We focus mainly on visuospatial WM, because this type of WM is most heavily studied across these different research fields. We also consider the possibility that improved spatial coding in attention and WM is a key component behind training and transfer. Next, we describe the neuroimaging, behavioural, pharmacological and electrophysiological evidence concerning the types of plasticity that mediate the changes in WM capacity, in both humans and non-human primates.

An improved knowledge of the mechanisms underlying changes in WM capacity could potentially be of practical use in future interventions. Moreover, because WM training is an experimentally well-controlled environmental manipulation, this research sheds light on the mechanisms through which the environment affects cognitive ability. Understanding these mechanisms may be relevant for understanding childhood development and how activities in daily life affect cognitive abilities.

Neural basis of working memory

Electrophysiological evidence from non-human primates — as well as some supporting evidence from rodent studies (BOX 1) — has made it possible to understand the basis of WM at the single-neuron level and to construct detailed computational models of WM that can account for variation in WM performance and capacity. In turn, these results have enabled the interpretation of human imaging studies, which reveal activation at the brain network level.

Working memory in non-human primates. The cellular basis of WM has been investigated extensively with electrophysiological recordings in macaques while they remember the spatial location or identity of a stimulus over a delay period lasting for a few seconds (typically 1–5 s). Such investigations reveal that single neurons in the PFC exhibit discharges that persist even after the physical stimuli that elicited them are no longer present[6,22] (known as persistent activity, or delay activity; FIG. 1a,b). Individual neurons exhibit persistent activity that is tuned for specific attributes of a remembered stimulus, such as its location, direction of motion or identity[23,24], or more abstract information about, for example, the category to which it belongs or its relevance in the context of a task[25,26]. Persistent activity during the delay period of WM tasks is thought to be generated by reverberating discharges in a network of interconnected neurons in the PFC (FIG. 1a), and between the PFC and other brain areas, such as the posterior parietal and inferior temporal cortex, and subcortical structures, including the basal ganglia and the mediodorsal nucleus of the thalamus[27,28].

Persistent activity in the PFC does not merely reflect properties of remembered stimuli but is predictive of WM performance. Behavioural trials in which persistent activity is diminished are associated with errors in recall[22,29]. Indeed, a near-linear relationship between persistent activity and the ability to remember the correct stimulus has been revealed in some WM tasks[30]. Levels of persistent activity are predictive of the accuracy of eye movements towards the remembered locations of visual stimuli, on a trial-to-trial basis[31]. Variations in WM performance attributable to a host of other phenomena (for example, the reduction of WM performance with ageing or owing to dual-task interference) are associated with different levels of persistent activity in the PFC[32,33]. In cases in which multiple brain areas display persistent activity, the monkeys' recall is most strongly tied to the levels of persistent activity in the PFC, as revealed by choice probability analysis[34].

Non-human primate studies that involve simple WM tasks, such as requiring detection of a change in stimulus colour of one item in an array of items after a brief delay, have obtained estimates of capacity that are similar to (though generally lower than) those obtained in humans[35–37]. Activity recorded during these tasks has revealed that the information that can be extracted from PFC neuron activity quickly reaches its maximum when more than one item is being remembered[35]. An interesting observation is that this capacity limitation seems to be independent for the two hemifields of vision[35], mirroring human results that suggest relatively faster decay of WM performance for multiple items presented in the same hemifield as opposed to along the vertical meridian[38]. The results suggest that within-hemifield competition limits the ability to maintain WM for multiple objects.

Computational models of working memory. Based on results from neurophysiological experiments of WM, biophysically realistic models of networks of neurons have been created in which persistent discharges are mediated by activity reverberating in the network through recurrent connections between neurons[39]. A 'ring' model has been particularly fruitful in simulating persistent activity relating to visuospatial WM: each node in the network represents a neuron with the peak of its receptive field at a different location in a zone of the visual field at a constant distance from the fovea, forming a ring[40] (FIG. 1c). The strength of connections between neurons in the model depends on the distance between their preferred locations in the visual field, in accordance with experimental

Choice probability
Probability that the firing rate of one neuron in two identical stimulus conditions is different depending on the subject's choice.

Receptive field
Area of the visual field where a stimulus will elicit the firing of neuron under study.

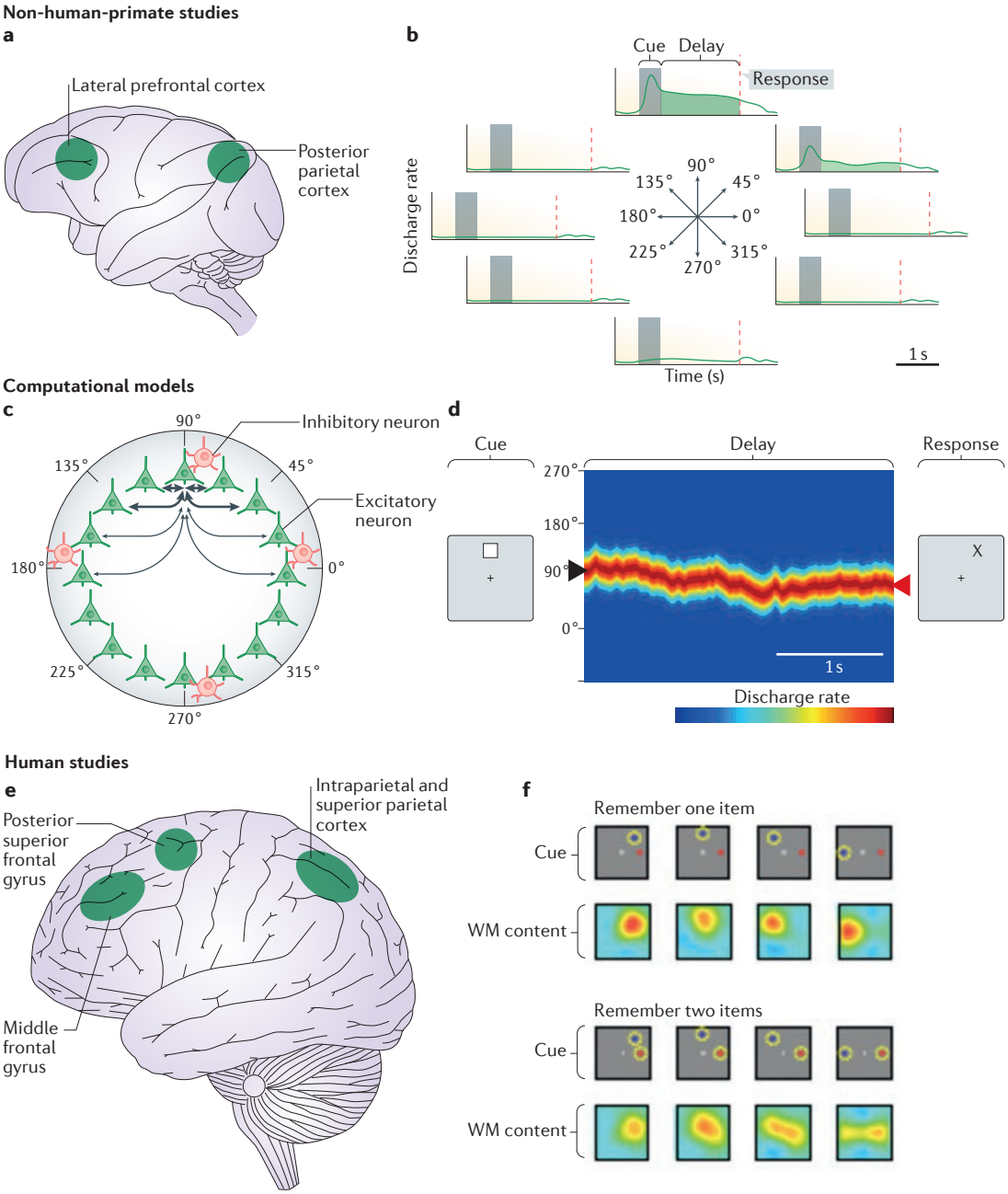

Figure 1 | **Neural basis of working memory. a** | Diagram of the monkey brain, with the areas implicated in visuospatial working memory (WM)[28] — the posterior parietal cortex and the lateral prefrontal cortex — highlighted in green. **b** | Schematic diagram of activity of a single prefrontal neuron during execution of the oculomotor delayed response task[22]. In this task, a monkey is presented with a cue located in a particular area of the visual field. The cue is then removed, and the monkey has to wait several seconds (the delay period) before moving its gaze to the location where the cue appeared. In this example, the neuron is activated during the presentation of the cue that is located at 90° relative to the horizon and continues to discharge during the delay period, in a spatially selective manner. **c** | Computational modelling proposes possible mechanisms for how information is retained in WM. Here, information can be coded in a cell-specific manner in excitatory neurons (triangles). Inhibitory cells are symbolized by circles. Activity is maintained through recurrent excitation[40]. Outputs of inhibitory neurons are distributed to all neurons in the network (not shown). **d** | Activity of the neural network, in which the colour represents the level of activity (blue: low activity; red: high activity). The bump of activity predicts the remembered location[31]. In this example, the bump 'drifts' such that the recalled location of the cue is ~30° away from the true location of the presented cue. **e** | Key areas of the human brain activated in visuospatial WM tasks and tasks that require spatially selective attention, including the intraparietal and superior parietal cortex, the posterior part of the superior frontal gyrus and the middle frontal gyrus[60–63]. **f** | The content of WM can be reconstructed from the functional MRI signal in the human parietal cortex during the delay period of a WM task in which subjects remember one or two items at different locations (top row). The lower row shows the images reconstructed from brain activity[74]. Part **f** is adapted with permission from REF. 74, Elsevier.

results: a pair of neurons that represent locations that are further apart on the ring are connected by a weaker connection than are neurons representing adjacent parts of the ring[40]. A cue briefly presented in the periphery of a ring at 90° relative to the horizontal axis is thus remembered by persistent activity of the '90° neuron' and its most strongly connected neighbouring cells.

Simulations of such a network of recurrent connections reveal that activity in the network behaves as a continuous attractor; the activity peak, or 'bump', that represents the firing rate of the population of neurons survives the disappearance of the stimulus, but the location it represents drifts slowly in time from that of the original location of the stimulus[31] (FIG. 1d). In this framework, the drift of the bump results in a loss in precision: greater drifts represent larger errors between the location represented by the persistent activity relative to the location represented by neuron activity during stimulus exposure. Items can also be forgotten because overall inhibition in the network decreases the firing frequency (that is, eliminates the bumps), such that it is not high enough to maintain recurrent activity. Experimental evidence confirms key predictions of this model. For example, the location that monkeys recall in a visuospatial WM task deviates precisely in the direction of the drift of neuronal activity recorded from single neurons at the end of the delay period[31]. The variability of discharges from single neurons and the level of correlation between neurons recorded simultaneously also vary, consistent with a drifting bump model[31] and contrary to the idea that inaccuracy in visuospatial WM is primarily due to the slow decaying of the bump of neuronal activity itself.

Mechanisms of maintaining items in WM that do not depend on persistent activity have also been proposed. Synaptic (as opposed to spiking) mechanisms are likely to mediate effects such as repetition suppression, in which a lower level of activity is observed for a previously presented stimulus[41]. Some models of WM therefore rely on synaptic processes, such as changes in presynaptic calcium levels, or fast Hebbian synaptic plasticity to mediate WM[42–44]. Repetition suppression, however, is limited to memory for repeated stimuli and could not play a part in other memory tasks such as those requiring a response towards a previously presented stimulus or reporting the identity of a remembered stimulus. Another alternative mechanism through which information about the properties of a preceding stimulus may be represented is through the dynamic pattern of activation of different neurons at each time point during a WM task[45]. Such a dynamic code does not strictly require persistent activity, but it is not incompatible with it either[46,47]. Information about stimulus properties may also be present in the magnitude, frequency and phase of rhythmic activity during the WM interval[48–50]. These factors may be related to the rhythmicity detected in EEG recordings during WM and to changes after WM training (reviewed below). WM is thus represented in the precise timing of discharges in models of biophysically realistic neurons with oscillatory activity[51]. An increase in WM load in such models results in increases of both theta and gamma frequencies, consistent with human data. Models of WM that are based on persistent activity still provide the most accurate and detailed account of neural and behavioural results obtained in neurophysiological experiments, particularly in visuospatial WM tasks[31]. A more detailed argument about the relative merits of persistent-activity-dependent models of WM versus competing models can be found elsewhere[52].

Computational models of working memory capacity. Models that propose that the activity representing each of the stored items held in memory acts as a bump attractor offer mechanistic understanding of capacity in spatial WM[53–55]. As more items are added to memory, the total population activity increases, as does the amount of activity of inhibitory interneurons in the system. However, when the capacity of the system is exceeded, the persistent activity representing some of the stimuli decays because of the increased inhibition, and these stimuli cannot be recalled at the end of the delay period. As a consequence, the model predicts that population activity reaches its maximum at a set point that corresponds to the maximum behavioural capacity. The PFC may represent stimulus content in WM, but another crucial function of the PFC is to provide a top-down signal that insulates stimulus-coding networks both in the PFC and in other areas that generate persistent activity, such as the posterior parietal cortex, from the mutually inhibitory effects of having multiple items stored in WM[53].

Various other mechanisms for the maintenance of multiple items have been proposed by alternative WM models that do not rely on persistent activity. The concurrent activity of a population of neurons coding for several multi-featured items in a pattern that allows extraction of stimulus properties may serve to store their attributes in memory[56]. Alternatively, a gain mechanism may represent different attributes of objects: that is, single neurons may exhibit tuning for one attribute, and their overall level of activity may be modulated by a second attribute[57]. The attributes of different items may yet be represented by the specific timing (phase) of spiking in a rhythmic code[51]. The capacity of WM in such a model would thus be determined by the number of items that can be simultaneously represented at different phases of rhythmic activity.

Computational models have identified crucial factors on which WM performance depends. These include the strength of synaptic interactions within and between stimulus-encoding networks, the strength of synaptic activity from a region providing excitatory input to a stimulus-encoding network (for example, from the PFC to the parietal cortex), the extent of synaptic arborization of individual neurons, the ability of individual neurons to exhibit bistability and the strength of NMDA receptor and dopamine D1 receptor (D1R) activation[53,54,58]. Higher strengths of synaptic connectivity are generally associated with higher ability of the network to sustain activity, particularly in the face of distraction. Increased density of NMDA receptors and increased D1R neurotransmission further stabilize network activity. Additionally, persistent activity with higher firing frequency, regardless of the particular mechanism that serves to enhance it, is associated with increased WM stability and higher WM

Repetition suppression
A phenomenon whereby a stimulus that has been repeated elicits a smaller response than does a stimulus that appears for the first time.

Bump attractor
A stable state of activation of a network, with a spatial location maximally activated and adjacent locations activated to a lesser extent, forming a bump.

Gain mechanism
Mechanism of representing of two variables, whereby the activity of a neuron depends linearly on one continuous variable, multiplied by second variable.

Tuning curve
A graph of firing-rate intensity depending on the location (or other dimension) of a stimulus.

capacity, because activity decays more easily in networks with lower discharge rates. Changes in these factors have been confirmed in empirical studies of WM training in human and animal models (discussed below). The same factors have also been implicated in the developmental maturation of WM capacity from childhood to adulthood and inter-individual variability in WM capacity among adults[59]. These models thus provide a mechanistic understanding of WM, and also suggest which changes could improve WM capacity.

Working memory and capacity in the human brain. Performance of WM tasks is associated with activity in a wide range of areas across the brain, depending on the type of stimulus. But there are also commonalities in the regions activated during different WM tasks; in particular, there is activity in a fronto-parietal network of regions, with the frontal regions including the middle frontal gyrus, the inferior frontal gyrus and the caudal part of the superior frontal sulcus (close to, or involving the frontal-eye field)[60,61] (FIG. 1e). The intraparietal cortex and caudal superior frontal gyrus are particularly strongly activated in visuospatial WM tasks[60] and in tasks that require spatially selective attention[62,63]. These shared areas that are activated during these different tasks could thus reflect the involvement of a spatial map of prioritized locations that is commonly used for visuospatial WM, spatially selective attention and a wide range of other cognitive tasks, including some verbal WM tasks[60,61].

Repeated sampling in fMRI allows tracking of the time course of the blood oxygenation level-dependent (BOLD) contrast signal, which is used as an indirect index of brain activity. In humans, there is sustained activity during the delays of WM tasks, consistent with the persistent delay activity found in non-human primates[64–66]. Techniques sensitive to the electric discharge of neurons, such as magnetoencephalography (MEG) and EEG, have revealed that during the delay in WM tasks, the activity of frontal regions becomes more synchronous with that in parietal regions[67], and that the degree of this synchronicity increases with the amount of information successfully kept in WM. Frontal and parietal regions also exhibit local increases in high-frequency oscillatory activity during WM tasks[68], a finding that is consistent with oscillatory models of WM[51].

Complicating the interpretation of increased BOLD activity during the delay are the observations that although the content of WM can be decoded from the delay activity, the BOLD signal used for this decoding is, in some conditions, independent of the sustained increase in BOLD activity during the delay, at least when decoding non-spatial stimulus information kept in WM — for example, stimulus orientation[69] or direction of motion[70]. This does not preclude the notion that spatial information is retained by some pattern of sustained or oscillatory firing during the delay. Both increases and decreases of the BOLD signal can be informative about the stimulus maintained in WM, reflecting excitatory and suppressive responses to the orientation and motion of stimuli. Moreover, the BOLD signal not only reflects firing frequency but is also influenced by, and possibly dominated by, sustained neuromodulatory input (BOX 2). Resolving these questions will require more information about the relationship of the BOLD signal to firing rate, the inhibitory–excitatory balance in neuronal activity and the microstructure of the PFC at the scale of the fMRI voxel, possibly through intracranial EEG recordings[71].

Another question is whether the retinotopic organization of neurons that exhibit delay activity during visuospatial WM task intervals, as found in macaque studies and modelled in computational studies, can also be found in humans. At least a coarse retinotopy has been observed in humans in both the parietal cortex and posterior PFC during performance of WM tasks, but not in more anterior parts of the PFC[72,73]. Consequently, the spatial location of the items that are kept in WM can be reconstructed from the pattern of activation measured during the delay period in either parietal regions or posterior frontal regions[74] (FIG. 1f). This points towards a consistency between computational models, neurophysiological data from non-human primates and human data, and suggests a standard model for the maintenance of spatial information in WM (shown in FIG. 1).

More-anterior parts of the PFC lack clear retinotopy, which could indicate that they provide top-down excitatory input to more posterior regions, without actually storing sensory information[53,75]. Alternatively, these

Box 2 | **Interpreting changes in the BOLD signal in working memory training**

With functional MRI (fMRI), brain activation can be evaluated by measuring the blood oxygenation level-dependent (BOLD) contrast signal. The signal reflects a change in the ratio of oxygenated to deoxygenated haemoglobin that occurs with brain activation and the increases in local blood volume. Concurrent electrophysiological and fMRI measurements show that the BOLD signal is related to both single-unit activity and local field potentials, but it responds to input and intracortical processing rather than pyramidal-cell output activity[71,130]. It is also possible that neuromodulation by noradrenaline, dopamine or acetylcholine could dominate the BOLD response[131].

A decrease in the BOLD signal in a certain area is often interpreted as an increase in 'efficiency' of the area in performing its function. However, the mechanisms of improving such efficiency are unclear. In the case of perceptual learning, one possible mechanism underlying improved efficiency is an increase in the specificity of stimulus coding, as shown in studies of monkeys trained to detect stimuli that were gradually degraded[86]. A narrower tuning curve would mean that the activity of fewer prefrontal or parietal neurons would be required to code a stimulus, and thus predict a smaller BOLD signal after training. Neurophysiological studies of working memory training in monkeys, however, reveal that training is associated with an increase in prefrontal pyramidal-cell firing rate and a flattening of tuning curves[84]. Another mechanism resulting in higher efficiency could be a decrease in the number of neurons that perform a cognitive process — in this case, maintenance of a certain amount of information. But, to our knowledge, there is still no clear evidence from neurophysiological studies of such efficiency-related mechanisms occurring in the cognitive domain.

An important caveat when interpreting BOLD-signal changes is that the signal directly reflects the amount of time that a particular brain area is activated, which could be indirectly reflected in reaction time — that is, the time from stimulus presentation to response. If training leads to a 10% higher firing rate in the 5% of neurons that code the stimulus, and the BOLD signal is linearly related to average firing rate, we would expect a 0.5% increase in the BOLD signal with training. If the reaction time, and thereby the length of time that the stimulus is kept in working memory, simultaneously decreases by 15%, any effect of the firing rate on the BOLD signal is drowned out or reversed. Unfortunately, changes in reaction times are rarely reported. Without information about the exact length of time that a certain area is active, we cannot say anything about the efficiency of the computations being performed.

regions could store information, albeit at an extremely small spatial scale. Neurophysiological studies in macaques suggest that the entire visual hemifield might be represented in PFC modules of an area no larger than 0.5×0.5 mm[76], thus explaining why fMRI decoding techniques — which tend to use much larger voxel sizes — may have failed in many studies to extract visuospatial information from the PFC of human brains during WM tasks.

Delay activity in the human parietal cortex (as assessed by measuring the BOLD signal in these regions) during delays in WM tasks is positively associated with trial-by-trial performance and inter-individual differences in WM capacity[77,78]. There is also a close relationship between EEG-detected delay activity in the parietal cortex and inter-individual differences in WM capacity in humans[79]: larger differences in polarity during the WM delay relative to baseline were associated with higher capacity.

Functional connectivity between cortical regions can be estimated by quantifying the covariance of activity between different brain regions during resting-state fMRI. A measure of the overall integration of connectivity into separate networks was found to be positively correlated with WM capacity in healthy adults[80]. Moreover, the synchronicity at several frequencies between frontal, parietal and visual areas was reported to positively correlate with inter-individual differences in WM capacity[67], as was the power of local oscillations, with a positive association most clearly seen for the gamma band (30–40 Hz)[68].

Working memory training

The brain exhibits remarkable plasticity, which enables it to adapt to environment changes, including increases in sensory stimulation, sensory deprivation or environment enrichment. Plasticity also enables the encoding of memories of episodes and places, and the learning of new habits and motor skills. The mechanisms underlying plasticity include the activation or suppression of genes (for example, of genes encoding receptors), changes in the strength of synapses, changes in the number of dendritic spines, changes in the arborization of axons and changes in myelination (for an overview, see REFS 81,82).

Several mechanisms are thus at hand, and can be initiated within minutes, days or weeks in response to changes in the environment. The question is: what evidence is there for the involvement of these plasticity mechanisms in the behavioural effects of WM training? Does cognitive training induce any of the same kinds of plasticity that occur during motor learning, declarative memory or perceptual priming?

Working memory training in non-human primates. Neurophysiological experiments have documented how PFC neurons in untrained monkeys respond to stimuli, and what changes occur to PFC neuron responses while and after the animals are trained to perform a task that requires WM for these stimuli (a process that takes ~0.5–1 years in non-human primates)[46,83–85]. The findings of these experiments are in line with predictions of WM models and results of human imaging studies. Training in WM tasks leads to an increase in the percentage of PFC neurons that are activated by the stimuli, an even greater increase in the percentage of neurons that exhibit persistent activity after the offset of the stimuli (that is, in the delay period of the task), and an increase in the mean firing rate of the activated neurons. Moreover, much larger populations of ventrolateral PFC neurons are active during the delay period of the task after training than before training[83,84]. These changes represent a selective and gradual increase of neuronal involvement in the task, rather than a recruitment of an area that was previously silent during the WM interval. By contrast, the proportion of dorsolateral PFC neurons that are activated during the presentation of stimuli changes little with training.

Despite the increased activity, the mean selectivity of individual neurons for trained stimuli diminishes after training, as shown by a decrease in the average neuronal selectivity for a stimulus, a decrease in the ability of a neuron to discriminate between stimuli and decreases in the amount of information that could be decoded from recordings from the entire PFC population[83,84]. That the selectivity of neurons for the trained stimuli decreases with training may initially be unexpected, given that monkeys were specifically trained to recognize and discriminate between stimuli, and given that an increase in stimulus selectivity has been observed in the PFC after perceptual training to recognize objects with added noise[86]. However, an important point to consider is that training in the WM task involved the initial presentation of the simple geometric stimuli at distinct locations, such that the stimulus set was familiar and highly discriminable before being incorporated in a WM task, rather than consisting of novel stimuli that the animals were trained to discriminate. Thus, a probable explanation of the decrease in neuronal selectivity for the trained stimuli is that a large portion of neuronal activity after training represents factors that are not specific to the stimuli, but that are related to the rules and execution of the WM task, such that the activity that is selective for the trained stimuli declines as a percentage of total neuronal activity[84].

In addition to changes in mean firing rate, training in a WM task is associated with decreased variability of firing rate across trials (as quantified by the Fano factor)[87]. Similarly, the level of across-trial correlation of neuronal discharges (known as spike-count correlation or noise correlation) also declines after training[88]. The effects of reduced variability and correlation suggest that training results in improved representation of stimulus properties in the population of PFC neurons through mechanisms that are independent of firing rate and that may be invisible to functional imaging methodologies.

Examining neural responses in monkeys to different stimuli before and after training, and quantifying the information that could be decoded with a classifier reveals that task-specific information is only weakly represented in the untrained PFC, whereas stimulus-specific information (for example, about the location and identity of the stimulus itself) can be reliably decoded[46]. By contrast, with training, a small population of PFC neurons becomes highly selective (at a

Functional connectivity
The likelihood that activity in one area leads to activity in the other.

Fano factor
Variance of spike counts divided by their mean, per unit of time.

Classifier
A model or algorithm achieving a classification decision of what a stimulus is, based on the combined activity of multiple neurons or voxels.

single-neuron level) for the task[46] and, at the same time, a larger population of neurons exhibits a modest degree of task-relevant information. Furthermore, after training, task-relevant information is present in the activity of neurons that are selective for shape information, often manifesting itself in relatively short time windows, in the midst of other large firing rate modulations that occurred during a trial. In essence, after training, PFC neurons multiplex different types of information, some of which were present before training (about the location and features of remembered stimuli) and some of which only appeared with training (representing task variables)[46].

Whereas these studies focused primarily on changes in neural activity between the completely naive stage and after full mastery of the task, other experiments provided comparisons between early and late phases of training — while performance continued to improve, in direct correspondence with human studies of training. Most importantly, delay-period activity increased in the PFC of monkeys during the course of training, as performance improved[84]. In the future, it would be interesting to evaluate the behavioural effects of WM training for a wider range of non-trained cognitive tasks, such as spatial attention, decision making and problem solving.

Working memory training in humans. Most neuroimaging studies of the effects of human WM training have focused on documenting changes in brain activity during performance of WM tasks, before and after training[89–103] (see Supplementary information S1 (table) for a full overview of available human neuroimaging studies of the functional effects of WM training). One question that has been addressed is to what extent training leads to new areas being activated, or changes the activation of areas that were already activated before WM training. The most consistent finding here is that WM training is associated with quantitative changes in the activity of networks that are already activated during the WM task before training. This is consistent with electrophysiological findings in monkeys of gradual increases in the activity of neurons during the WM delay, as performance in the task improved with training[83]. It is also in line with the resulting behavioural changes, which do not constitute the learning of a new behaviour, but rather a gradual improvement by around 0.6 standard deviations of a capacity from its baseline.

Both increases and decreases in task-related BOLD activity in different regions have been found to be associated with increases in WM capacity. An increase could reflect a higher firing rate during the delay or an increase in the number of cells with delay activity, as suggested by electrophysiological recordings in monkeys[83,84]. Higher-frequency firing is also associated with higher capacity, or more-stable representations, in computational studies[53,55,59]. A similar change in activation is furthermore observed during development of WM from childhood into adulthood (BOX 3).

Decreases have often been interpreted as reflecting improvements in efficiency, although the mechanisms of this increase in efficiency are not clear. Interpretations of changes in BOLD signals are complicated by the fact that the relationship between BOLD activity and firing rate is not straightforward (BOX 2). The problems in linking the firing rate to the BOLD signal and the dependence of changes in behaviour during scanning therefore limit the conclusions that can be drawn regarding underlying mechanisms. A cautious interpretation is thus that these fMRI studies point to the areas of change but do not inform us about the underlying cellular mechanisms.

The most consistently implicated loci of WM-training-related changes are the association areas of the PFC and the parietal cortex, including the network of regions that are known to be associated with WM capacity (FIG. 1e). If this network is used by a wider range of WM tasks[60,61,104], as well as in spatial attention[62,63], the robust association of WM training with functional changes in this network could explain the transfer effects of training. That is, the transfer effect may result from the fact that a wide range of tasks rely on partly the same network of brain regions, which change in response to training.

Certain regions, such as the medial temporal lobe, which is instrumental in declarative memory, do not seem to be involved in the training effect. Visual association areas, which have been implicated in perceptual priming, are also less often differentially activated after training. Together, these data indicate that WM training induces a specific type of learning, involving plasticity of higher association areas of the parietal (mainly intraparietal) cortex and PFC.

Increases in the activity of the striatum during WM tasks have also been associated with improvement during training, and striatal activity before training has been reported to predict the amount of transfer after training[91,93]. One study using repeated training found that WM-related activity in the striatum follows a nonlinear change during training, with an initial increase followed by a later decrease[99]. Although striatal changes are not found as consistently as those implicating the frontal and parietal cortex in WM-training-related changes, they are also predictive of change in WM capacity during development (BOX 3). A possible role of the striatum is further suggested by the associations between training improvement and the dopamine system — particularly D2Rs, which are abundant in the striatum (see 'Working memory training and dopamine' section).

Working memory training and connectivity. Changes in connectivity between neurons constitute the main cellular mechanism of plasticity. It is therefore logical to look for evidence of such changes in the imaging data as well. In one study[101], adults were imaged at rest, before and after 6 weeks of practicing verbal WM tasks that required manipulation of the material held in WM. Connectivity in the fronto-parietal network increased after training, and this change in connectivity positively correlated with improved performance of the trained WM tasks. Another investigation of resting-state connectivity (this time assessed with arterial spin labelling) found mainly increased connectivity between medial cortical regions, and smaller increases in a fronto-parietal network[102]. In one study, changes in functional connectivity were compared in one group of adults performing n-back WM

Box 3 | **Development of working memory during childhood**

Working memory (WM) capacity increases throughout childhood and adolescence, and studying the neural changes associated with this development is one way to learn about the mechanisms of WM capacity. Structural connectivity, as measured by magnetic resonance signal in white matter, has been related to WM capacity in 12-month-old infants[132], during childhood development[133–135] and during ageing[136]. Studies of the structural and functional changes associated with this development have found that higher capacity is associated with higher blood oxygenation level-dependent activity in the intraparietal, superior frontal and dorsolateral frontal regions[134,137–139]. Developmental studies of children and adolescents have indicated that striatal activity during WM task performance predicts future changes in WM capacity, which could imply that the striatum is involved in the development of WM capacity[134,140,141]. Computational models predict that the strength of connectivity between the prefrontal cortex and its cortical afferents is a crucial determinant of capacity, and that this is also likely to be an important factor during the normal course of development of WM from childhood to adulthood[53]. Data from adolescent and adult macaque monkeys also show that overall increases in the strength of synaptic connections between prefrontal-cortex neurons are associated with the development of WM[111].

training (FIG. 2), with a control group training on tracking of multiple visual objects[105]. Only in the group performing the WM task was there an increase in functional connectivity. The increase was detected between the parietal cortex and a superior frontal region, as well as between the parietal cortex and the dorsolateral PFC (FIG. 1e), and the increase in connectivity was significantly correlated with training-related improvement in performance on the trained tasks.

In another study[106], transcranial magnetic stimulation was applied over the parietal cortex, and the electrical response induced in other regions — that is, regions functionally connected to the parietal cortex — was mapped. One group performed WM training over 24 days and was compared to a control group who were trained on the computer game *Tetris*. WM training increased functional connectivity between the parietal and occipital areas, and between parietal and frontal areas (Brodmann area 4 (BA4), BA8 and BA9) specifically in trials with higher WM loads. This improvement in functional connectivity was also associated with the amount of improvement in non-trained cognitive tasks. Because the connectivity to frontal regions only increased in trials with higher loads, the authors interpreted the findings as being reflective not of permanent, structural changes, but rather of changes in oscillatory synchrony, possibly driven by changes in the balance between excitatory and inhibitory transmission[68]. An alternative interpretation is that there were structural changes in connections that are specifically recruited for higher loads. For example, a combined computational modelling and fMRI study suggested that a top-down, excitatory input from the PFC to the parietal cortex is specifically recruited at higher WM loads[53].

Further evidence of changes in connectivity comes from a study that used MEG to measure connectivity at rest, before and after training[107]. Children aged 8–11 years performed 5 weeks of mainly visuospatial WM training. Connectivity between a fronto-parietal network and the lateral occipital cortex as well as within parts of the parietal cortex increased in the training group and was positively correlated with improved WM capacity.

The mechanisms underlying changes in functional connectivity could be a stronger synaptic connectivity between neurons or an activity-dependent increase in the myelination of the connecting axons[108,109]. As reviewed above, measures of white-matter volume and structure are associated with WM capacity, and there is indeed evidence of changes in increases in white-matter density in the parietal lobe after WM training[110].

In conclusion, several studies suggest that stronger connections between cortical regions, in particular frontal and parietal regions, play a key part in improving WM with training. Stronger functional connectivity among frontal regions has been linked to better performance of WM tasks in monkeys[111], although it has not been specifically investigated in training studies. The observed effects of WM training on connectivity in humans is also in line with computational models that showed that strengthening functional connectivity from the PFC to the parietal cortex during the delay period enhances the stability and capacity of the local networks that maintain spatial representations[53,59].

Working memory training and dopamine. Dopamine is one of the most well-studied neurotransmitters implicated in WM and has also been associated with WM training. One study investigated the effect of WM training on dopamine-receptor density by scanning participants with PET before and after WM training using predominately visuospatial WM tasks[112]. In both scans, the density of D1Rs and D2Rs was mapped during rest with SCH 23390 (a D1R antagonist) and raclopride (a D2R antagonist), respectively. Changes in WM capacity were associated with changes in cortical D1R density: overall, there was a decrease in density, but this was best characterized as an inverted-U function in which the amount of change in receptor density was dependent on baseline receptor density. This result is consistent with the inverted-U graph of the effects of cortical application of D1R antagonists on delay activity in macaques[113] and with computational studies suggesting a role of dopamine in improving the ability to maintain information in WM despite distractions[55,58]. The changes in D1R density in humans can also be related to changes in mice after WM training[114]: such training resulted in an increase in animals' performance on a wide range of cognitive tasks compared to the control group, and enhanced sensitivity to D1R-mediated neuronal activation in the prelimbic cortex.

Another PET study in humans focused instead on the dynamics of dopamine release during performance of a WM task. D2R occupancy was measured using raclopride before and after WM training[115]. Relative to controls, the WM-training group showed an increase in release of dopamine in the caudate while performing the task.

The dopamine system is also implicated in the effects of WM training by a number of genetic studies. Two studies[116,117] found that variability in, or close to, the gene encoding dopamine transporter 1 (*DAT1*; also known as *SLC6A3*), the presynaptic transporter of dopamine into the cell, can explain some variability in how much individuals improve during training. *DAT1*

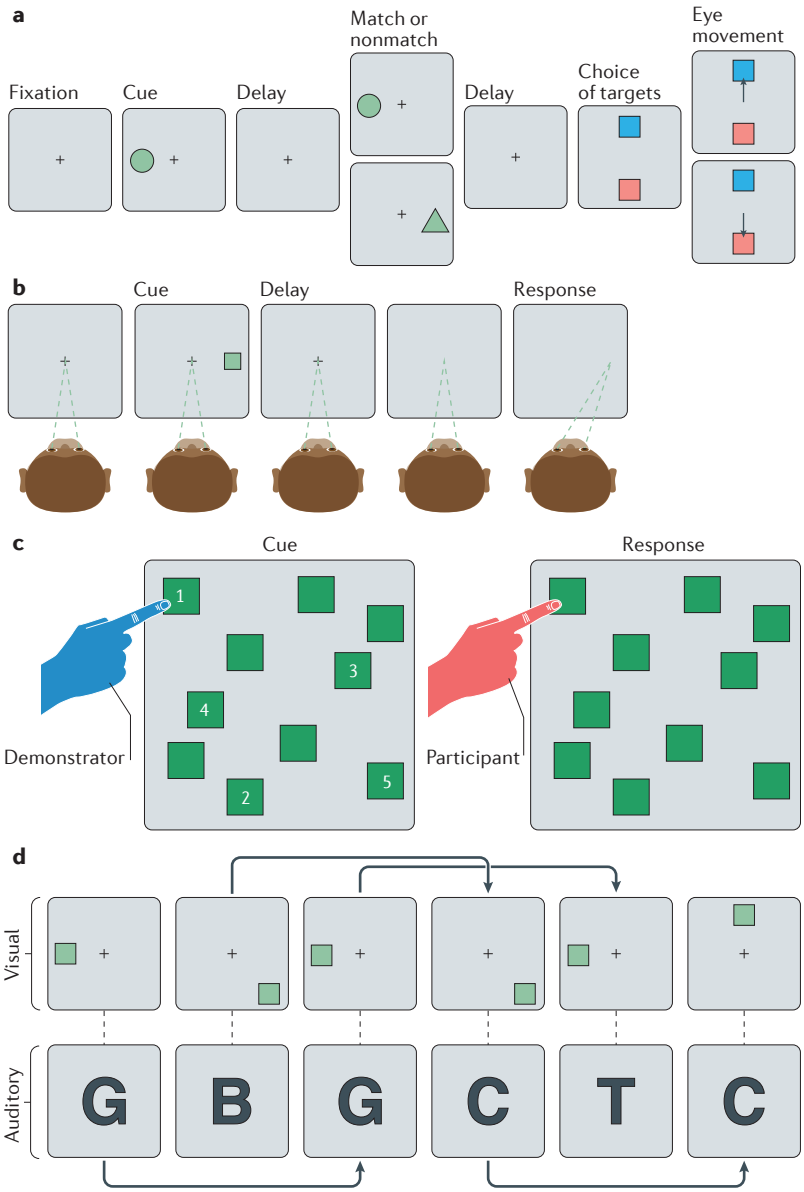

Figure 2 | **Working memory tasks. a** | An example of a match–nonmatch working memory (WM) task that can be used to evaluate the effects of training in monkeys[83]. The monkey fixates on a central point and observes two stimuli presented in sequence, separated by delay periods. The monkey has then to select a blue target if the two stimuli were identical, or a red target if they differed. **b** | A sequence of frames illustrates the oculomotor delayed response task. The monkey fixates on a central point and then, after cue presentation, must continue fixating until the fixation point disappears, at which point it must move its gaze towards the location of the previously presented cue (as in FIG. 1). **c** | The span-board task, a visuospatial WM task used in human studies. The experimenter points to a sequence of boxes. The positions and orders are kept in WM, and then the sequence is repeated by the test subject. **d** | A dual two-back task. The participant is asked to press a button if the position of the square matches its position two frames earlier, or if the letter aurally presented matches one presented two letters earlier.

variability has also been associated with the generally lower WM capacity in children with ADHD[118]. In a study of children and adolescents, a single-nucleotide polymorphism (rs1800497) in the coding region of ankyrin repeat and protein kinase domain-containing protein 1 (*ANKK1*) was identified as being associated with relative improvement on trained tasks after WM training[119]. The rs1800497 *ANKK1* polymorphism that was associated with a larger effect of training has previously been associated with a lower density of dopamine D2Rs in the striatum[120]. Interestingly, the same polymorphism interacts with BOLD activity in the ventral striatum, such that carriers show a stronger relationship between BOLD activity in the ventral striatum and WM capacity than do non-carriers[121]. Thus, fMRI, PET and genetic studies suggest that dopamine and the striatum play an important part in relative improvement with WM training.

Conclusions and future directions

Human neuroimaging studies found the most-consistent changes in brain activity after WM training to be related to activity-association areas of the frontal and parietal lobes — regions that are known to be key for WM capacity (FIG. 3). Several studies suggest that stronger inter-areal connectivity and increased synchronicity between the frontal and parietal lobes are associated with inter-individual differences in WM capacity, and that this connectivity is strengthened by training. These results are also consistent with findings from monkey electrophysiological studies that suggest that changes in activity of the PFC are related to more cells expressing delay activity, higher overall levels of activity during WM task performance and changes in connectivity within the region. Computational studies suggest that stronger functional connectivity — for example, between the PFC and the parietal cortex — leads to a higher firing rate in the parietal networks that code spatial information during the delay period of WM tasks. The models also suggest that, together, a higher firing rate and stronger connectivity result in a greater stability of the networks coding the information held in WM.

The transfer from one trained task to a non-trained task is probably related to a shared reliance of such tasks on a fronto-parietal network. The function of this network has been suggested to be 'cognitive control', but it could also simply be to code salient spatial locations used for spatially selective attention, visuospatial WM and cognitive processes that depend on spatial coding. This could be a general function for human cognition, and the computational studies provide a mechanistic model for this (FIG. 1).

The dopaminergic system is implicated in the effects of WM training in several ways. Genetic studies suggest that variability in dopaminergic signalling-related genes may contribute to inter-individual differences in improvements after training. PET studies in humans, and studies of rodents, show that WM training leads to changes in dopamine release and dopamine receptor density. The implications of dopamine could partly be related to the role of the striatum, which does not necessarily retain information during WM trials but could facilitate cortical plasticity.

Notably, there are no consistent changes associated with WM training in the primary visual cortex (as in perceptual learning) or in the medial temporal lobe (as in declarative memory). The importance of the striatum

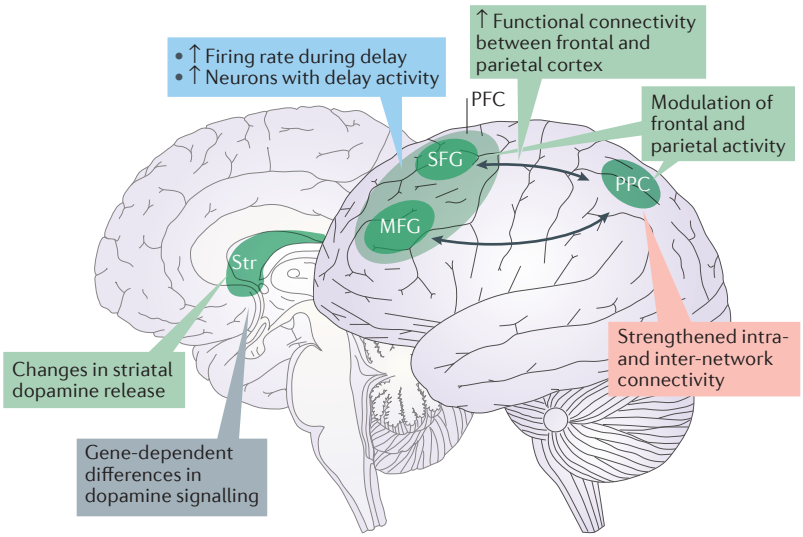

Figure 3 | **Summary of factors underlying training-induced increases in capacity.** Electrophysiological studies in monkeys show that training of working memory (WM) leads to a larger number of prefrontal cortex (PFC) neurons with delay activity and higher firing frequency during the delay period. Studies using functional MRI (fMRI) and electroencephalography in humans associate stronger fronto-parietal functional connectivity[53,67,80] and structural connectivity[133,135,141] with higher capacity, and training studies show increased fronto-parietal connectivity after WM training[101,105–107]. Functional connectivity could be related to either myelination or stronger synaptic connectivity. fMRI in humans has consistently shown changes (both increases and decreases) in brain activity in frontal and parietal regions after training (see Supplementary information S1 (table)). Simulations using neuronal networks suggest that the capacity of a stimulus-encoding network can be increased by stronger functional connectivity between the PFC and parietal cortex[53,59]. Genetic studies have associated greater training improvement with certain polymorphisms of the postsynaptic dopamine D2 receptor gene (*DRD2*)[119] and the presynaptic sodium-dependent dopamine transporter gene (*DAT1*)[116,117], both of which are highly expressed in the striatum (Str). fMRI in humans suggests that the Str mediates the effects of WM training. Positron-emission tomography studies have detected changes in striatal dopamine release[115] and cortical dopamine D1 receptor density[112]. MFG, middle frontal gyrus; PPC, posterior parietal cortex; SFG, superior frontal gyrus.

indicates that there may be some similarities with motor learning and implicit memory, although motor regions are not involved in WM training. The hallmark of cognitive plasticity is suggested to be the changes in connectivity of higher association areas in the frontal and parietal lobes.

Further research into the plasticity of WM functions will require research about the concept of WM itself. In this Review, we describe a model of visuospatial WM and persistent activity, with information coded in the spatial tuning of neurons. There are probably several different neuronal mechanisms that could contribute to short-term retention of information, including short-term synaptic memory and oscillatory information representation. It is also possible that the mechanisms for maintenance of memories for non-spatial stimuli rely on different mechanisms. Future nomenclatures of memory based on different types of neuronal mechanisms will hopefully be developed.

In human WM-training studies, there are still gaps in the knowledge about how to interpret and integrate findings of changes in the BOLD signal and changes in the pattern of neural activity during the delay, and whether these represent mechanisms of capacity. Comparing fMRI with intracranial EEG recordings during performance of the same visuospatial WM task could fill parts of this gap. There is also a lack of electrophysiological data from non-human primates on inter-regional connectivity and how this affects firing rate, local networks and capacity. Recordings from the striatum of rodents or non-human primates over the course of WM training would provide important information.

One of the questions to be addressed in the future is whether plasticity during cognitive training could be enhanced. This could be relevant in, for example, individuals with learning disabilities, in whom plasticity is decreased. The implication of the dopamine system in several WM training studies suggests that pharmacological interventions affecting this system could be a promising place to start.

Another line of research has suggested that stimulation of the brain with weak electric currents, so-called transcranial direct current stimulation (TDCS), could enhance both performance and learning. Indeed, TDCS over the dorsolateral PFC enhanced improvement in WM training in one study[122]. There is still a lack of knowledge about the effects of TDCS. Moreover, the method has a low spatial resolution, and many results are contradictory. Nevertheless, in theory, TDCS has the potential to translate knowledge from neuroimaging studies to interventions.

Finally, relying on neuroscience findings to inform psychology and educational sciences about previously unknown behavioural or psychological mechanisms of learning continues to be a challenge. Understanding more about the neural mechanisms of cognitive plasticity, its time course and facilitating factors could, in the future, optimize behavioural paradigms for training, and lead to more-effective interventions.

1. Miller, G. A., Galanter, E. & Pribram, K. H. *Plans and the Structure of Behavior* (Rinehart and Winston, 1960).
2. Baddeley, A. D. & Hitch, G. J. in *The Psychology of Learning and Motivation: Advances in Research and Theory* (ed. Bower, G. A.) 47–90 (Academic Press, 1974).
3. Pribram, K. H., Mishkin, M., Rosvold, H. E. & Kaplan, S. J. Effects on delayed-response performance of lesions of dorsolateral and ventromedial fronal cortex of baboons. *J. Comp. Physiol. Psychol.* **45**, 565–575 (1952).
4. Pribram, K. H., Ahumada, A., Hartog, J. & Ross, L. in *The Frontal Granular Cortex and Behavior* (eds Warren, J. M. & Akert, K.) 65 (McGraw-Hill, 1964).
5. Kubota, K. & Niki, H. Prefrontal cortical unit activity and delayed alternation performance in monkey. *J. Neurophysiol.* **34**, 337–347 (1971).
6. Fuster, J. M. & Alexander, G. E. Neuron activity related to short-term memory. *Science* **173**, 652–654 (1971).
7. Goldman-Rakic, P. S. in *Handbook of Physiology* (ed. Mountcastle, V. B.) 373–417 (1987).
8. Cowan, N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* **24**, 87–185 (2001).
9. Kane, M. J. *et al.* For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. *Psychol. Sci.* **18**, 614–621 (2007).
10. Kyllonen, P. C. & Christal, R. E. Reasoning ability is (little more than) working-memory capacity?! *Intelligence* **14**, 389–433 (1990).
11. Gathercole, S. E., Brown, L. H. & Pickering, S. J. Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. *Educat. Psychol.* **70**, 177–194 (2003).
12. Klingberg, T., Forssberg, H. & Westerberg, H. Training of working memory in children with ADHD. *J. Clin. Exp. Neuropsychol.* **24**, 781–791 (2002).
13. Klingberg, T. *et al.* Computerized training of working memory in children with ADHD — a randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* **44**, 177–186 (2005).

14. Jaeggi, S. M., Buschkuehl, M., Jonides, J. & Perrig, W. J. Improving fluid intelligence with training on working memory. *Proc. Natl Acad. Sci. USA* **105**, 6829–6833 (2008).
15. Cortese, S. *et al.* Cognitive training for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *J. Am. Acad. Child Adolesc. Psychiatry* **54**, 164–174 (2015).
16. Schwaighofer, M., Fischer, F. & Buhner, M. Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educat. Psychol.* **50**, 138–166 (2015).
17. Peijnenborgh, J. C., Hurks, P. M., Aldenkamp, A. P., Vles, J. S. & Hendriksen, J. G. Efficacy of working memory training in children and adolescents with learning disabilities: a review study and meta-analysis. *Neuropsychol. Rehabil.* **17**, 1–28 (2015).
18. Brehmer, Y., Westerberg, H. & Backman, L. Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Front. Hum. Neurosci.* **6**, 63 (2012).
19. Green, C. T. *et al.* Will working memory training generalize to improve off-task behavior in children with attention-deficit/hyperactivity disorder? *Neurotherapeutics* **9**, 639–648 (2012).
20. Bigorra, A., Garolera, M., Guijarro, S. & Hervas, A. Long-term far-transfer effects of working memory training in children with ADHD: a randomized controlled trial. *Eur. Child Adolesc. Psychiatry* http://dx.doi.org/10.1007/s00787-015-0804-3 (2015).
21. Conklin, H. M. *et al.* Computerized cognitive training for amelioration of cognitive late effects among childhood cancer survivors: a randomized controlled trial. *J. Clin. Oncol.* **33**, 3894–3902 (2015).
22. Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* **61**, 331–349 (1989).
This classic paper offers a description of how PFC activity represents spatial information of remembered stimuli, thereby providing a neural code for spatial WM.
23. Goldman-Rakic, P. S. Cellular basis of working memory. *Neuron* **14**, 477–485 (1995).
24. Constantinidis, C. & Wang, X. J. A neural circuit basis for spatial working memory. *Neuroscientist* **10**, 553–565 (2004).
25. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291**, 312–316 (2001).
26. Wallis, J. D., Anderson, K. C. & Miller, E. K. Single neurons in prefrontal cortex encode abstract rules. *Nature* **411**, 953–956 (2001).
27. Pasternak, T. & Greenlee, M. W. Working memory in primate sensory systems. *Nat. Rev. Neurosci.* **6**, 97–107 (2005).
28. Constantinidis, C. & Procyk, E. The primate working memory networks. *Cogn. Affect. Behav. Neurosci.* **4**, 444–465 (2004).
29. Zhou, X. *et al.* Working memory performance and neural activity in the prefrontal cortex of peri-pubertal monkeys. *J. Neurophysiol.* **110**, 2648–2660 (2013).
30. Constantinidis, C., Franowicz, M. N. & Goldman-Rakic, P. S. The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nat. Neurosci.* **4**, 311–316 (2001).
31. Wimmer, K., Nykamp, D. Q., Constantinidis, C. & Compte, A. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat. Neurosci.* **17**, 431–439 (2014).
This study demonstrates that persistent activity in the delay period of a spatial WM task can be modelled as a bump attractor and determines performance in the task.
32. Watanabe, K. & Funahashi, S. Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. *Nat. Neurosci.* **17**, 601–611 (2014).
33. Wang, M. *et al.* Neuronal basis of age-related working memory decline. *Nature* **476**, 210–213 (2011).
34. Mendoza-Halliday, D., Torres, S. & Martinez-Trujillo, J. C. Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat. Neurosci.* **17**, 1255–1262 (2014).
This paper illustrates that persistent activity is absent from the visual middle temporal area but emerges for the first time in the medial superior temporal area, and is further transmitted to the PFC.
35. Buschman, T. J., Siegel, M., Roy, J. E. & Miller, E. K. Neural substrates of cognitive capacity limitations. *Proc. Natl Acad. Sci. USA* **108**, 11252–11255 (2011).
36. Heyselaar, E., Johnston, K. & Pare, M. A change detection approach to study visual working memory of the macaque monkey. *J. Vis.* **11**, 11 (2011).
37. Lara, A. H. & Wallis, J. D. Capacity and precision in an animal model of visual short-term memory. *J. Vis.* **12**, 13 (2012).
38. Stormer, V. S., Alvarez, G. A. & Cavanagh, P. Within-hemifield competition in early visual areas limits the ability to track multiple objects with attention. *J. Neurosci.* **34**, 11526–11533 (2014).
39. Wang, X. J. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* **24**, 455–463 (2001).
40. Compte, A., Brunel, N., Goldman-Rakic, P. S. & Wang, X. J. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* **10**, 910–923 (2000).
This paper documents a biophysically plausible computational model of persistent activity that has been instrumental for the study of spatial WM.
41. Grill-Spector, K., Henson, R. & Martin, A. Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* **10**, 14–23 (2006).
42. Mongillo, G., Barak, O. & Tsodyks, M. Synaptic theory of working memory. *Science* **319**, 1543–1546 (2008).
The paper provides a computational model that does not depend on persistent activity and can account for some memory phenomena, based on synaptic mechanisms.
43. Sugase-Miyamoto, Y., Liu, Z., Wiener, M. C., Optican, L. M. & Richmond, B. J. Short-term memory trace in rapidly adapting synapses of inferior temporal cortex. *PLoS Comput. Biol.* **4**, e1000073 (2008).
44. Sandberg, A., Tegner, J. & Lansner, A. A working memory model based on fast Hebbian learning. *Network* **14**, 789–802 (2003).
45. Stokes, M. G. *et al.* Dynamic coding for cognitive control in prefrontal cortex. *Neuron* **78**, 364–375 (2013).
46. Meyers, E. M., Qi, X. L. & Constantinidis, C. Incorporation of new information into prefrontal cortical activity after learning working memory tasks. *Proc. Natl Acad. Sci. USA* **109**, 4651–4656 (2012).
47. Crowe, D. A., Averbeck, B. B. & Chafee, M. V. Neural ensemble decoding reveals a correlate of viewer- to object-centered spatial transformation in monkey parietal cortex. *J. Neurosci.* **28**, 5218–5228 (2008).
48. Salazar, R. F., Dotson, N. M., Bressler, S. L. & Gray, C. M. Content-specific fronto-parietal synchronization during visual working memory. *Science* **338**, 1097–1100 (2012).
49. Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D. & Miller, E. K. Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron* **76**, 838–846 (2012).
50. Siegel, M., Warden, M. R. & Miller, E. K. Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl Acad. Sci. USA* **106**, 21341–21346 (2009).
51. Lundqvist, M., Herman, P. & Lansner, A. Theta and gamma power increases and alpha/beta power decreases with memory load in an attractor network model. *J. Cogn. Neurosci.* **23**, 3008–3020 (2011).
52. Riley, M. R. & Constantinidis, C. Role of prefrontal persistent activity in working memory. *Front. Syst. Neurosci.* **9**, 181 (2016).
53. Edin, F. *et al.* Mechanism for top-down control of working memory capacity. *Proc. Natl Acad. Sci. USA* **106**, 6802–6807 (2009).
This study integrates computational modelling and fMRI to test a model of top-down attention and WM storage.
54. Macoveanu, J., Klingberg, T. & Tegner, J. Neuronal firing rates account for distractor effects on mnemonic accuracy in a visuo-spatial working memory task. *Biol. Cybern.* **96**, 407–419 (2007).
55. Durstewitz, D., Seamans, J. K. & Sejnowski, T. J. Neurocomputational models of working memory. *Nat. Neurosci.* **3** (Suppl.), 1184–1191 (2000).
56. Matthey, L., Bays, P. M. & Dayan, P. A probabilistic palimpsest model of visual short-term memory. *PLoS Comput. Biol.* **11**, e1004003 (2015).
57. Botvinick, M. & Watanabe, T. From numerosity to ordinal rank: a gain-field model of serial order representation in cortical working memory. *J. Neurosci.* **27**, 8636–8642 (2007).
58. Durstewitz, D., Seamans, J. K. & Sejnowski, T. J. Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* **83**, 1733–1750 (2000).
59. Edin, F., Macoveanu, J., Olesen, P., Tegner, J. & Klingberg, T. Stronger synaptic connectivity as a mechanism behind development of working memory-related brain activity during childhood. *J. Cogn. Neurosci.* **19**, 750–760 (2007).
60. Rottschy, C. *et al.* Modelling neural correlates of working memory: a coordinate-based meta-analysis. *Neuroimage* **60**, 830–846 (2012).
61. Nee, D. E. *et al.* A meta-analysis of executive components of working memory. *Cereb. Cortex* **23**, 264–282 (2013).
62. Jerde, T. A., Merriam, E. P., Riggall, A. C., Hedges, J. H. & Curtis, C. E. Prioritized maps of space in human frontoparietal cortex. *J. Neurosci.* **32**, 17382–17390 (2012).
63. Ikkai, A. & Curtis, C. E. Common neural mechanisms supporting spatial working memory, attention and motor intention. *Neuropsychologia* **49**, 1428–1434 (2011).
64. Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S. & Passingham, R. E. The prefrontal cortex: response selection or maintenance within working memory? *Science* **288**, 1656–1660 (2000).
65. Curtis, C. E., Rao, V. Y. & D'Esposito, M. Maintenance of spatial and motor codes during oculomotor delayed response tasks. *J. Neurosci.* **24**, 3944–3952 (2004).
66. Courtney, S. M., Ungerleider, L. G., Keil, K. & Haxby, J. V. Transient and sustained activity in a distributed neural system for human working memory. *Nature* **386**, 608–611 (1997).
67. Palva, J. M., Monto, S., Kulashekhar, S. & Palva, S. Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc. Natl Acad. Sci. USA* **107**, 7580–7585 (2010).
68. Roux, F. & Uhlhaas, P. J. Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct WM information? *Trends Cogn. Sci.* **18**, 16–25 (2014).
69. Ester, E. F., Sprague, T. C. & Serences, J. T. Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* **87**, 893–905 (2015).
This paper demonstrates that attributes of remembered stimuli can be decoded from fMRI activity of the PFC in humans, reaffirming neurophysiological findings that placed the maintenance of information in the frontal lobe rather than the sensory cortices.
70. Emrich, S. M., Riggall, A. C., Larocque, J. J. & Postle, B. R. Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *J. Neurosci.* **33**, 6516–6523 (2013).
71. Goense, J. B. & Logothetis, N. K. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr. Biol.* **18**, 631–640 (2008).
72. Sereno, M. I., Pitzalis, S. & Martinez, A. Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science* **294**, 1350–1354 (2001).
73. Silver, M. A. & Kastner, S. Topographic maps in human frontal and parietal cortex. *Trends Cogn. Sci.* **13**, 488–495 (2009).
74. Sprague, T. C., Ester, E. F. & Serences, J. T. Reconstructions of information in visual spatial working memory degrade with memory load. *Curr. Biol.* **24**, 2174–2180 (2014).
75. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
76. Constantinidis, C., Franowicz, M. N. & Goldman-Rakic, P. S. Coding specificity in cortical microcircuits: a multiple electrode analysis of primate prefrontal cortex. *J. Neurosci.* **21**, 3646–3655 (2001).
77. Todd, J. J. & Marois, R. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* **428**, 751–754 (2004).
78. Xu, Y. & Chun, M. M. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* **440**, 91–95 (2006).
79. Vogel, E. K. & Machizawa, M. G. Neural activity predicts individual differences in visual working memory capacity. *Nature* **428**, 748–751 (2004).
80. Stevens, A. A., Tappon, S. C., Garg, A. & Fair, D. A. Functional brain network modularity captures inter- and intra-individual variation in working memory capacity. *PLoS ONE* **7**, e30468 (2012).
81. Holtmaat, A. & Svoboda, K. Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat. Rev. Neurosci.* **10**, 647–658 (2009).

82. van Praag, H., Kempermann, G. & Gage, F. H. Neural consequences of environmental enrichment. *Nat. Rev. Neurosci.* **1**, 191–198 (2000).
83. Meyer, T., Qi, X. L., Stanford, T. R. & Constantinidis, C. Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J. Neurosci.* **31**, 6266–6276 (2011).
84. Qi, X. L., Meyer, T., Stanford, T. R. & Constantinidis, C. Changes in prefrontal neuronal activity after learning to perform a spatial working memory task. *Cereb. Cortex* **21**, 2722–2732 (2011).
References 83 and 84 document effects of WM training at the level of single neurons in the PFC of non-human primates.
85. Qi, X. L., Meyer, T., Stanford, T. R. & Constantinidis, C. Neural correlates of a decision variable before learning to perform a Match/Nonmatch task. *J. Neurosci.* **32**, 6161–6169 (2012).
86. Rainer, G. & Miller, E. K. Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* **27**, 179–189 (2000).
87. Qi, X. L. & Constantinidis, C. Variability of prefrontal neuronal discharges before and after training in a working memory task. *PLoS ONE* **7**, e41053 (2012).
88. Qi, X. L. & Constantinidis, C. Correlated discharges in the primate prefrontal cortex before and after working memory training *Eur. J. Neurosci.* **36**, 3538–3548 (2012).
89. Garavan, H., Kelley, D., Rosen, A., Rao, S. M. & Stein, E. A. Practice-related functional activation changes in a working memory task. *Microsc. Res. Tech.* **51**, 54–63 (2000).
90. Hempel, A. *et al.* Plasticity of cortical activation related to working memory during training. *Am. J. Psychiatry* **161**, 745–747 (2004).
91. Olesen, P. J., Westerberg, H. & Klingberg, T. Increased prefrontal and parietal brain activity after training of working memory. *Nat. Neurosci.* **7**, 75–79 (2004).
92. Westerberg, H. *et al.* Computerized working memory training after stroke — a pilot study. *Brain Inj.* **21**, 21–29 (2007).
93. Dahlin, E., Neely, A. S., Larsson, A., Backman, L. & Nyberg, L. Transfer of learning after updating training mediated by the striatum. *Science* **320**, 1510–1512 (2008).
94. Jolles, D. D., Grol, M. J., Van Buchem, M. A., Rombouts, S. A. & Crone, E. A. Practice effects in the brain: changes in cerebral activation after working memory practice depend on task demands. *Neuroimage* **52**, 658–668 (2010).
95. Mozolic, J. L., Hayasaka, S. & Laurienti, P. J. A cognitive training intervention increases resting cerebral blood flow in healthy older adults. *Front. Hum. Neurosci.* **4**, 16 (2010).
96. Schneiders, J. A., Opitz, B., Krick, C. M. & Mecklinger, A. Separating intra-modal and across-modal training effects in visual working memory: an fMRI investigation. *Cereb. Cortex* **21**, 2555–2564 (2011).
97. Subramaniam, K. *et al.* Computerized cognitive training restores neural activity within the reality monitoring network in schizophrenia. *Neuron* **73**, 842–853 (2012).
98. Schneiders, J. A. *et al.* The impact of auditory working memory training on the fronto-parietal working memory network. *Front. Hum. Neurosci.* **6**, 173 (2012).
99. Kuhn, S. *et al.* The dynamics of change in striatal activity following updating training. *Hum. Brain Mapp.* **34**, 1530–1541 (2013).
100. Schweizer, S., Grahn, J., Hampshire, A., Mobbs, D. & Dalgleish, T. Training the emotional brain: improving affective control through emotional working memory training. *J. Neurosci.* **33**, 5301–5311 (2013).
101. Jolles, D. D., van Buchem, M. A., Crone, E. A. & Rombouts, S. A. Functional brain connectivity at rest changes after working memory training. *Hum. Brain Mapp.* **34**, 396–406 (2013).
102. Takeuchi, H. *et al.* Effects of working memory training on functional connectivity and cerebral blood flow during rest. *Cortex* **49**, 2106–2125 (2013).
103. Subramaniam, K. *et al.* Intensive cognitive training in schizophrenia enhances working memory and associated prefrontal cortical efficiency in a manner that drives long-term functional gains. *Neuroimage* **99**, 281–292 (2014).
104. Klingberg, T., Roland, P. E. & Kawashima, R. Activation of multi-modal cortical areas underlies short-term memory. *Eur. J. Neurosci.* **8**, 1965–1971 (1996).
105. Thompson, T. W., Waskom, M. L. & Gabrieli, J. D. E. Intensive working memory training produced functional changes in large-scale frontoparietal networks. *J. Cogn. Neurosci.* **28**, 575–588 (2016).
106. Kundu, B., Sutterer, D. W., Emrich, S. M. & Postle, B. R. Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *J. Neurosci.* **33**, 8705–8715 (2013).
This study uses transcranial magnetic stimulation and EEG to show increases in connectivity after WM training in children.
107. Astle, D. E., Barnes, J. J., Baker, K., Colclough, G. L. & Woolrich, M. W. Cognitive training enhances intrinsic brain connectivity in childhood. *J. Neurosci.* **35**, 6277–6283 (2015).
Using resting-state MEG, this study shows how training-induced improvements in WM capacity are associated with increases in functional connectivity.
108. Gibson, E. M. *et al.* Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. *Science* **344**, 1252304 (2014).
109. Yeung, M. S. *et al.* Dynamics of oligodendrocyte generation and myelination in the human brain. *Cell* **159**, 766–774 (2014).
110. Takeuchi, H. *et al.* Training of working memory impacts structural connectivity. *J. Neurosci.* **30**, 3297–3303 (2010).
111. Zhou, X. *et al.* Age-dependent changes in prefrontal intrinsic connectivity. *Proc. Natl Acad. Sci. USA* **111**, 3853–3858 (2014).
112. McNab, F. *et al.* Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science* **323**, 800–802 (2009).
Here, PET is used to map changes in cortical D1R binding associated with WM training.
113. Sawaguchi, T. & Goldman-Rakic, P. S. D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science* **251**, 947–950 (1991).
114. Wass, C. *et al.* Dopamine D1 sensitivity in the prefrontal cortex predicts general cognitive abilities and is modulated by working memory training. *Learn. Mem.* **20**, 617–627 (2013).
115. Backman, L. *et al.* Effects of working-memory training on striatal dopamine release. *Science* **333**, 718 (2011).
116. Brehmer, Y. *et al.* Working memory plasticity modulated by dopamine transporter genotype. *Neurosci. Lett.* **467**, 117–120 (2009).
117. Soderqvist, S. *et al.* Dopamine, working memory, and training induced plasticity: implications for developmental research. *Dev. Psychol.* **48**, 836–843 (2012).
118. Shang, C. Y. & Gau, S. S. Association between the *DAT1* gene and spatial working memory in attention deficit hyperactivity disorder. *Int. J. Neuropsychopharmacol.* **17**, 9–21 (2014).
119. Soderqvist, S., Matsson, H., Peyrard-Janvid, M., Kere, J. & Klingberg, T. Polymorphisms in the dopamine receptor 2 gene region influence improvements during working memory training in children and adolescents. *J. Cogn. Neurosci.* **26**, 54–62 (2014).
120. Hirvonen, M. *et al.* C957T polymorphism of the dopamine D2 receptor (*DRD2*) gene affects striatal DRD2 availability *in vivo*. *Mol. Psychiatry* **9**, 1060–1061 (2004).
121. Nymberg, C. *et al. DRD2/ANKK1* polymorphism modulates the effect of ventral striatal activation on working memory performance. *Neuropsychopharmacology* **39**, 2357–2365 (2014).
122. Richmond, L. L., Wolk, D., Chein, J. & Olson, I. R. Transcranial direct current stimulation enhances verbal working memory training performance over time and near transfer outcomes. *J. Cogn. Neurosci.* **26**, 2443–2454 (2014).
123. Delatour, B. & Gisquet-Verrier, P. Functional role of rat prelimbic–infralimbic cortices in spatial memory: evidence for their involvement in attention and behavioural flexibility. *Behav. Brain Res.* **109**, 113–128 (2000).
124. Erlich, J. C., Bialek, M. & Brody, C. D. A cortical substrate for memory-guided orienting in the rat. *Neuron* **72**, 330–343 (2011).
125. Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62–68 (2012).
126. Light, K. R. *et al.* Working memory training promotes general cognitive abilities in genetically heterogeneous mice. *Curr. Biol.* **20**, 777–782 (2010).
127. Cassanelli, P. M. *et al.* Working memory training triggers delayed chromatin remodeling in the mouse corticostriatothalamic circuit. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **60**, 93–103 (2015).
128. Ren, M., Cao, V., Ye, Y., Manji, H. K. & Wang, K. H. Arc regulates experience-dependent persistent firing patterns in frontal cortex. *J. Neurosci.* **34**, 6583–6595 (2014).
129. Yuan, Y., Mao, H. & Si, J. Cortical neural responses to previous trial outcome during learning of a directional choice task. *J. Neurophysiol.* **113**, 1963–1976 (2015).
130. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
131. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
132. Short, S. J. *et al.* Associations between white matter microstructure and infants' working memory. *Neuroimage* **64**, 156–166 (2013).
133. Vestergaard, M. *et al.* White matter microstructure in superior longitudinal fasciculus associated with spatial working memory performance in children. *J. Cogn. Neurosci.* **23**, 2135–2146 (2011).
134. Ullman, H., Almeida, R. & Klingberg, T. Structural maturation and brain activity predict future working memory capacity during childhood development. *J. Neurosci.* **34**, 1592–1598 (2014).
135. Ostby, Y., Tamnes, C. K., Fjell, A. M. & Walhovd, K. B. Morphometry and connectivity of the fronto-parietal verbal working memory network in development. *Neuropsychologia* **49**, 3854–3862 (2011).
136. Kennedy, K. M. & Raz, N. Aging white matter and cognition: differential effects of regional variations in diffusion properties on memory, executive functions, and speed. *Neuropsychologia* **47**, 916–927 (2009).
137. Klingberg, T., Forssberg, H. & Westerberg, H. Increased brain activity in frontal and parietal cortex underlies the development of visuo-spatial working memory capacity during childhood. *J. Cogn. Neurosci.* **14**, 1–10 (2002).
138. Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L. & Bunge, S. A. Neurocognitive development of the ability to manipulate information in working memory. *Proc. Natl Acad. Sci. USA* **103**, 9315–9320 (2006).
139. Scherf, K. S., Sweeney, J. A. & Luna, B. Brain basis of developmental change in visuospatial working memory. *J. Cogn. Neurosci.* **18**, 1045–1058 (2006).
140. Klingberg, T. Childhood cognitive development as a skill. *Trends Cogn. Sci.* **18**, 573–579 (2014).
141. Darki, F. & Klingberg, T. The role of fronto-parietal and fronto-striatal networks in the development of working memory: a longitudinal study. *Cereb. Cortex* **25**, 1587–1595 (2015).

Acknowledgements
This work was supported by US National Institutes of Health grant R01EY017077 and The Tab Williams Family Endowment (to C.C.).

Competing interests statement
The authors declare no competing interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (table)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF