CellPress

Review
The Distributed Nature of Working Memory

Thomas B. Christophel,[1,2,3,12,]* P. Christiaan Klink,[4,5,6,12]
Bernhard Spitzer,[7] Pieter R. Roelfsema,[5,6,8,12] and
John-Dylan Haynes[1,2,3,9,10,11,12]

Studies in humans and non-human primates have provided evidence for storage of working memory contents in multiple regions ranging from sensory to parietal and prefrontal cortex. We discuss potential explanations for these distributed representations: (i) features in sensory regions versus prefrontal cortex differ in the level of abstractness and generalizability; and (ii) features in prefrontal cortex reflect representations that are transformed for guidance of upcoming behavioral actions. We propose that the propensity to produce persistent activity is a general feature of cortical networks. Future studies may have to shift focus from asking where working memory can be observed in the brain to how a range of specialized brain areas together transform sensory information into a delayed behavioral response.

Storage of Working Memory Contents

Cognition critically depends on the ability to memorize information and change it adaptively. This ability is typically ascribed to working memory, which has often been conceptualized as comprising two basic functions: (i) short-term storage of information and (ii) executive processes that retain no information themselves but control what is retained [1,2]. Here, we argue that instead of being discrete, these functions could be considered extreme functional specializations of a distributed working memory network combining storage and control. Different brain regions might contribute to working memory in accordance with the general nature of their representations. Thus, we consider working memory to rely on the interplay between brain regions that retain low-level or more abstract sensory information, and the transformation of this information for guidance of memory-guided behavioral plans.

Traditionally, working memory research in neuroscience has focused on the question of where contents are stored across delay periods. However, this is subject to ongoing debates [3–8]. Current views emphasize the representation of memorized contents in the prefrontal [3,8–10], parietal [7], or sensory [5,11–13] cortices. Some views more clearly consider the representation of memorized contents at multiple levels of the cortical hierarchy [2,4,14–17]. Localizing working memory storage has proven difficult. The first neural models of working memory were primarily shaped by lesion studies in humans and non-human primates [18–20]. For example, a seminal study [18] performed lesions in prefrontal cortex (PFC) of monkeys and found deficits in their ability to maintain task-relevant information across brief delays, while performance remained intact for tasks that did not require memorization. This and similar work [21–25] led to the belief that working memory is primarily a function of prefrontal cortex.

Later, electrophysiological [26] and functional magnetic resonance imaging (fMRI) studies (reviewed in [27]) revealed persistent neural activity in various brain regions during memory delays, including sensory cortices [28]. However, if a brain region X (say PFC) exhibits delay

[1]Bernstein Center for Computational Neuroscience, Charité Universitätsmedizin, Berlin, Germany
[2]Berlin Center for Advanced Neuroimaging, Charité Universitätsmedizin, Berlin, Germany
[3]Clinic for Neurology, Charité Universitätsmedizin, Berlin, Germany
[4]Department of Neuromodulation & Behaviour, Netherlands Institute for Neuroscience, Amsterdam, the Netherlands
[5]Department of Vision & Cognition, Netherlands Institute for Neuroscience, Amsterdam, the Netherlands
[6]Department of Psychiatry, Academic Medical Center, University of

CrossMark

period activity during working memory delays, and if lesions to X decrease memory performance, it does not necessarily mean that X encodes the working memory content during that delay. Instead, region X might just have an influence on storage taking place in other regions. Any neural signal that encodes working memory content should exhibit at least the following two properties. First, it should contain information about what is memorized, meaning that different memory contents lead to different patterns of activity (stimulus selectivity). Second, stimulus-selectivity activity should be present over extended delays in the absence of the stimulus (persistent activity or delay-period activity). The current review focuses on observations of such **persistent stimulus-selective activity** (see Glossary) in single-cell and local field potential recordings in non-human primates and in human neuroimaging [electroencephalography (EEG) and fMRI]. Please note that delay activity in both sensory and prefrontal regions occasionally fails to extend throughout the maintenance period [29], or can be absent altogether [30,31]. For this reason, working memory mechanisms that do not rely on persistent activity (**activity-silent working memory** [6,32–34]), that postulate encoding in **dynamic firing trajectories** [35,36] or that involve gating of spiking activity by local field potentials [37] have also been postulated (see Outstanding Questions). Importantly, persistent activity does not imply that a single selective neuron is firing throughout the entire delay period. Instead, it means that the activity of a neural population encodes stimulus-specific information at any point during the delay. Furthermore, persistent stimulus-selective activity has been proposed to require recurrent excitation [38,39], changes in synaptic facilitation or combinations of both [32,34,40].

Electrophysiological Recordings in Non-Human Primates

First reports of persistent working memory signals date back to the 1970s and describe studies on non-human primates performing delayed response tasks. Most of these early studies focused on the PFC because they were inspired by previous findings that lesions in the PFC cause deficits in delayed response tasks [18]. In one study [41], monkeys were presented with two lamps, one in the left and one in the right visual hemifield. At the beginning of a trial, one of the lamps was briefly turned on. After a delay of a few seconds, the monkeys were trained to press a key on the left or right, corresponding to the memorized side of the light. Activity in a subset of cells in the dorsolateral prefrontal cortex (dlPFC) was modulated during the delay period. Some cells discriminated between left and right stimuli, thus potentially providing evidence for content-specific delay-period activity in the dlPFC. However, in this study the position of the lights was correlated with the location of the response keys so that the selective delay-period activity might as well have represented motor preparation rather than stimulus memory. The identification of which feature is being represented is a general problem with memory-related representations that is elaborated upon below.

Subsequent studies revealed cells in the PFC that encode memorized positions during the delay period, even when controlling for response preparation [42]. Content-specific delay-period activity in the PFC has been shown for many sensory features (Figure 1, left): objects and natural images [43–45], color [46], and visual motion [29]; also for more abstract features such as numerosity [47] and for other modalities such as the frequency of tactile vibrations [48]. Detailed analyses of visual motion signals in the PFC during working memory has furthermore revealed that they exhibit a tuning profile that is comparable to those found in sensory regions [29,49].

While working memory deficits after lesions in the PFC led most early electrophysiological studies to focus on the PFC, further recordings in non-human primates have also revealed persistent stimulus-selective activity for spatial location in V1 [50,51], for motion in MT [52] and MST [49], for shapes and real-life stimuli in V4 and the temporal cortex [53–55], and for color in the inferior temporal cortex [56,57] (Figure 1, left). Persistent stimulus-selective activity has also been observed in the sensory cortices of other modalities, such as for pure tones in the auditory cortex [58], for haptic texture in the primary somatosensory cortex [59] and, to some extent, for

Amsterdam, Amsterdam, The Netherlands
[7]Department of Experimental Psychology, University of Oxford, Oxford, UK
[8]Department of Integrative Neurophysiology, Centre for Neurogenomics and Cognitive Research, VU University, Amsterdam, The Netherlands
[9]Berlin School of Mind and Brain, Humboldt Universität, Berlin, Germany
[10]Cluster of Excellence NeuroCure, Charité Universitätsmedizin, Berlin, Germany
[11]Department of Psychology, Humboldt Universität zu Berlin, Berlin, Germany
[12]These authors contributed equally to this work.

*Correspondence:
tbchristophel@gmail.com
(T.B. Christophel,).

vibration in the secondary somatosensory cortex [60]. An overview of evidence for persistent stimulus-selective activity in primate electrophysiology is shown in Figure 1 (left; see Table S1 in the supplemental information online).

Apart from the PFC and sensory cortices, persistent stimulus-selective activity has furthermore been reported in the primate frontal eye fields (FEFs) [61], parietal cortex [62], premotor cortex [60,63] and the medial temporal lobe [64]. Activity preceding delayed responses has also been reported in subcortical areas such as the mediodorsal nucleus of the thalamus (see [65]), superior colliculi [66], basal ganglia [67], and spinal cord [68] (not included in Figure 1).

Multivariate Decoding in Humans

In the human brain, it can be more challenging to identify content-specific delay-period activity due to the limited spatial resolution of noninvasive neuroimaging techniques. Early neuroimaging studies did not distinguish between content-selective and non-selective delay-period signals [27]. Others reported category-specific differences, for example, between spatial locations and objects [69] or houses and faces [70]. However, with the development of multivariate decoding (also known as multivoxel pattern analysis; MVPA [71–74]), content-specific activity also became accessible to human neuroimaging studies. This was further advanced by the development of inverse encoding models [75–78]. Such studies have identified delay-period information about sensory features in a range of sensory cortical regions (Figure 1, right). Low-level visual features such as orientation, color, motion, or complex patterns can be decoded from early visual areas [79–83]. Auditory stimuli can be read out from the primary auditory cortex [84,85]. Delay-period information about complex visual pattern stimuli has also been found in parietal areas [81,86]. More recent human imaging studies eventually found content-specific delay activity also in frontal areas [77,85,87–92], which was not observed in early MVPA studies [81,82]. Retinotopically organized area FEF, for instance, carries information about memorized spatial position in several tasks [87]. Also, complex shapes [92] and oriented gratings [77] can be decoded from human FEF (often also referred to as superior precentral sulcus). Furthermore, delay-period representations of natural objects [88], Chinese script [93], auditory pure tones [85], and to a limited extend, oriented gratings [77] can be decoded from signals in the lateral PFC. Additional evidence for content-specific delay activity in human PFC comes from EEG studies [89–91] that showed tactile stimulus frequency representations in prefrontal electrode locations.

Taken together, findings in non-human primates and humans reveal that the content of working memory is widely distributed and can be found across sensory, parietal, temporal, and prefrontal cortices (see Figure 1 and Table S1 in the supplemental information online, for an overview; see also [77]).

Why Are Working Memory Representations Distributed?

The presence of working memory signals in both low-level and high-level cortices suggests some kind of distributed representation (Box 1). There are several accounts for this finding (Figure 2). For example, it could in theory reflect a full duplication of sensory information in the PFC (Figure 2A; [8]). While the redundancy of such duplicate representations might, at first sight, stand in contradiction to evolutionary frugality [4,5], it could be useful in establishing the robustness of working memory contents against distraction (see below and [43,94–97]). However, there also appear to be substantial differences in the functional roles of persistent stimulus-selective activity between low-level and high-level brain regions. In the following, we discuss two of these differences (see also Box 2): (i) persistent stimulus-selective activity reflects different levels of abstraction with early sensory regions encoding simple features of stimuli in **sensory representations** and prefrontal regions encoding memories using more **abstract representations** or a verbal format (Figure 2B); and (ii) persistent stimulus-selective activity reflects different functional roles of these representations across multiple cortical regions,

Glossary

Abstract representation: high-level representation of memorized stimuli that is less detailed with regards to specific physical stimulus features, but that allows for generalization across a large number of stimuli that share the same semantic property or that belong to the same category.
Activity-silent working memory: neural mechanisms of working memory that are not directly detectable as changes in spike rate (or BOLD activity), for example: temporary changes in synaptic efficacy.
**Distributed representations:** neural representations of a memorized content (i.e., stimulus information) that can be found in multiple neural populations (see Box 4).
Dynamic firing trajectories: encoding of working memory contents in a dynamic temporal pattern of spiking activity.
Persistent stimulus-selective activity: any neural activity that takes place after a sensory stimulus is no longer present and that holds information about that stimulus.
Sensory representation: low-level representation of memorized stimuli that carries detailed information about the originally presented stimulus.
Turing machine: hypothetical machine, invented by mathematician Alan Turing as a thought experiment on computational operations (see Box 3).
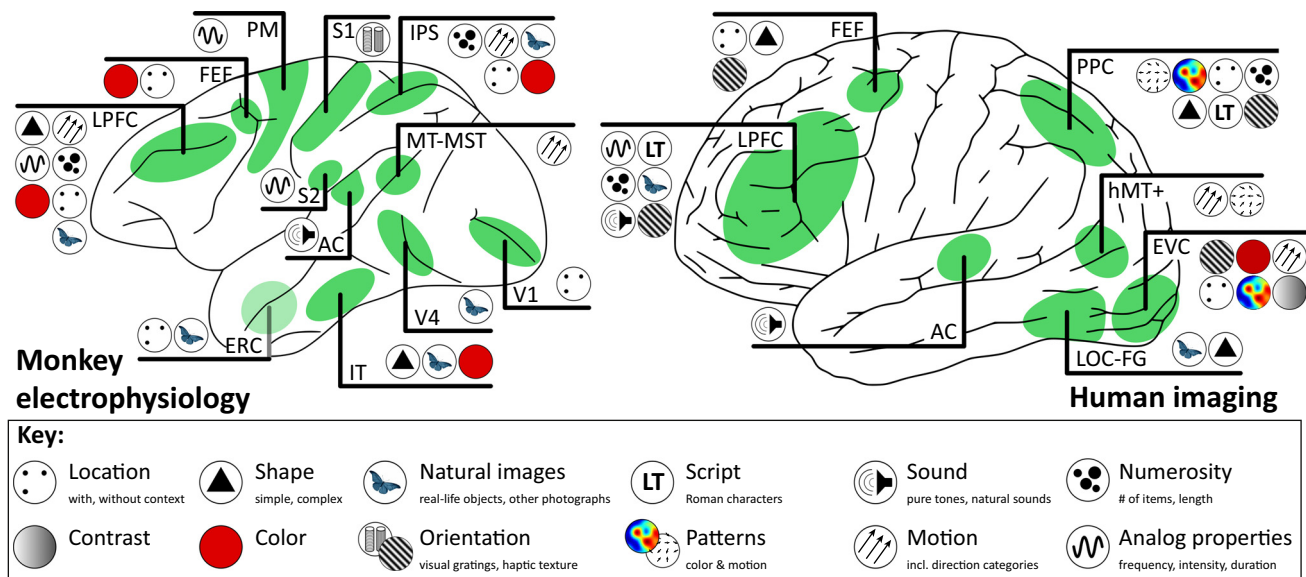
Figure 1. Overview of Content-Specific Activity during Working Memory Delays in the Macaque (Left) and Human (Right) Brain. Icons indicate persistent stimulus-selective activity for each stimulus type indicated by the icon (see legend) at the respective locations. Brain areas are identified by abbreviations (a full list of individual studies is reported in Table S1 in the supplemental information online). Both left- and right-sided effects are shown on the left hemisphere. Data from delay-periods during which subjects can prepare a specific motor output instead of memorizing a stimulus (e.g., delayed response tasks, e.g., [137]) and tasks that involve any form of explicit mental transformation [192,193] or explicitly learned association [132] are excluded. AC, auditory cortex; ERC, enthorinal cortex; EVC, early visual cortex; FEF, frontal eye fields; FG, fusiform gyrus; hMT+, human analog to MT/MST; IPS, intraparietal sulcus; IT, inferior temporal cortex; LOC, lateral occipital complex; lPFC, lateral prefrontal cortex; PM, premotor cortex; PPC, posterior parietal cortex.

ranging from representing the incoming low-level sensory features to planning a memory-dependent behavioral response (Figure 2C).

Different Levels of Abstraction

Imagine a task that involves memorizing one out of two visually presented animals (say a yellow lion and a grey elephant, see Figure 2). It is possible to memorize the distinction between these animals at different levels of abstraction, for example, using low-level visual features such as outlines or colors, or abstract categories like type-of-animal. Importantly, the memory representations can go beyond the features presented in the samples. In our example, it would be possible to memorize images of animals based on their typical sounds, their verbal labels, or even the emotions they evoke. In early work many authors believed working memory to be of a purely verbal nature [98–102], but we now know that working memories can also have basic sensory qualities [103–106]. Working memory signals in sensory and prefrontal cortices could reflect such different levels of abstraction. This specialization between sensory and prefrontal regions for different levels of abstraction is a general feature of cortical processing [2,107]. In the field of working memory, it is supported by several lines of evidence.

An early route to distinguishing sensory from abstract memory representations has been to require participants to memorize nonsense patterns or shapes that are hard to verbalize. Behavioral studies using this strategy have shown that working memory can rely exclusively on visual features [103]. Content-specific delay-period activity for such nonsense stimuli has been found in ventral occipitotemporal cortex, dorsal occipitoparietal areas, and the frontal eye fields, but not in PFC [53,81,86,92]. Interpreting these findings, however, relies on interpreting null results, which can be difficult for imaging data in particular.

> ### Box 1. Distributed Representations
>
> The term **distributed representation** is ambiguous and requires clarification. Several distinct interpretations can be identified.
>
> (1) Local pattern information: it has been repeatedly shown that the orientation of a grating stimulus during a working memory delay can be decoded from patterns of brain activity in primary visual cortex [79]. In this case a distributed representation means that the orientation information is disseminated across a local population of units within a single area.
>
> (2) Separable information in multiple brain regions: orientation information can also be decoded from signals in the lPFC, independent of the decodable information in the primary visual cortex [77]. Distributed here means that there are multiple local response patterns in parallel; each of which can be independently used to decode orientation. In this review, this is what is meant by the term distributed representation.
>
> (3) Inseparable information across multiple brain regions: finally, information might be encoded in global patterns of brain activity that is not encoded at the level of individual regions. In this case, a single area would not allow for decoding of such information, but considering the signals across multiple regions together would allow the extraction of information.
>
> Distinguishing between Cases 2 and 3 experimentally requires an assessment of whether the different regions contain sufficient information on their own to decode a stimulus feature, or whether only the combination of signals across areas allows decoding. Thus, an important criterion is whether the information in the multiple regions is redundant. The distinction between Case 2 and 3 is probably gradual rather than categorical. Memory codes could, for example, be redundant or synergistic so that the information present in two regions is less or more than the sum, respectively.
>
> If multiple, redundant, working memory representations indeed exist, it raises the interesting question which of them will be used to solve any given task. Lesion studies are one way to answer this question, as partially evidenced from the fact that many working memory tasks can be performed even after PFC lesions [20,156]. Another approach could be to investigate the choice probabilities [157] associated with activity in particular areas to analyze which signals are maximally predictive of behavioral choices in a memory task.

A different approach to dissociate levels of abstraction is to use stimuli with both low-level and abstract features and to vary the degree to which the task emphasizes sensory detail. One recent neuroimaging study [88] capitalized on this strategy using real-life visual objects. In one (visual) condition, participants were shown an object fragment and asked whether this fragment belonged to a memorized item or not. In another (nonvisual) condition, they had to report whether a new object (e.g., a watch) belonged to the same category as the memorized item (e.g., a clock). Sensory areas exhibited delay-period information about the stimulus during the visual task but not during the nonvisual task. Prefrontal areas showed the opposite pattern. Similarly, the PFC was found to represent color information when individuals were asked to discern clearly categorizable colors, but not when discriminating subtle color hues [31,46]. This distinction between memory representations for fine-grained color hues and color category is further supported by recent behavioral experiments and modeling work [108].

Another way of testing whether a brain region encodes abstract information is to assess to which degree neuronal memory representations generalize across different stimuli. In one series of studies [109,110], for instance, macaques were trained to memorize morphed images of animals whose low-level visual features could vary either within category or across a category boundary between cats and dogs. Prefrontal neurons generalized more across different exemplars of the same category than neurons in the inferior temporal cortex. Likewise, converging evidence from human [91] and non-human primate electrophysiology [63] indicates that neurons in frontal areas that were initially demonstrated to represent memorized tactile vibration frequencies [48], can in fact represent parametric memories from a range of sensory modalities (visual, auditory, and tactile). Frontal areas thus appear to encode these memories in an abstract, amodal magnitude format [111], while parametric delay activity in modality-specific sensory cortices is mostly absent [89,91,112]. Prefrontal areas have also been shown to host memory representations of abstract number information [47,113] in formats that may generalize across sensory modalities [114,115].

Taken together, these findings are compatible with a division of labor, in which sensory regions encode low-level details and prefrontal regions encode abstract, categorical information that
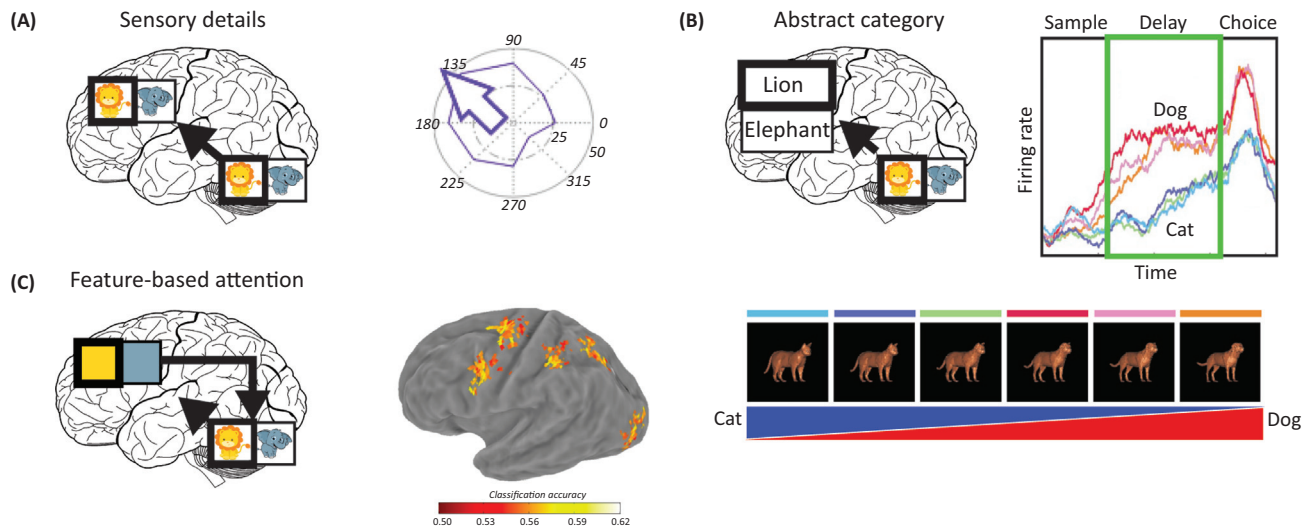
Figure 2. Potential Explanations for Persistent Stimulus-Selective Activity in Prefrontal Cortex during Working Memory Delays, Shown Here for the Example of Memorizing a Lion (versus an Elephant). (A) The prefrontal cortex maintains sensory details with similar resolution to sensory brain regions [29,77]. The right shows a plot of responses of a cell in monkey lateral prefrontal cortex to different directions of motion, which exhibits similar properties to responses in MT [194]. (B) Representation of abstract category information. The right panel shows responses to six morphs gradually varying between pure cat and pure dog images. This cell largely encodes the perceived category with the three morphs that are perceptually classified as dogs yielding a similar and high categorical response [109,195]. (C) Prefrontal cortex prospectively encodes feature-based attention during the delay that prepares for an upcoming search for the target. This can serve as a potential bias signal for feature-selective representations in early visual cortex [196].

generalizes across modalities [5,116,117]. A computational advantage of storing low-level features in sensory areas is that it circumvents the necessity to duplicate low-level feature spaces for persistent storage in higher brain regions (Figure 2A, [5,116]). Persistent stimulus-selective activity in the PFC can then be used to build flexible and task-dependent representations, such as arbitrary categorical labels (Figure 2B; [116]). However, some studies have suggested that also low-level features can be represented in the PFC [46,77], which appears to be at odds with a distinction based purely on the level of abstraction. It has been proposed that such representations of low-level sensory features in the PFC might reflect memorization using a nonsensory neural code [31] (see also Box 2). For example, orientation information can be memorized using symbolic, numerical cues and this does not necessarily involve persistent stimulus-selective in sensory cortices (see Supplementary Material in [79]). Alternatively, as we outline next, the prevalence of sensory information in the PFC might reflect the transformation of sensory information in a form that serves to prepare for an upcoming behavioral response.

Transformation from Sensory Input to a Behavioral Response

A central function of working memory is to bridge the gap between a stimulus and a response that is to be executed while the stimulus is no longer present. The PFC seems to encode the to-be-memorized information in a format that can directly guide behavior ([5,14,15,118,119]). Persistent activity thus not only reflects a retrospective memory of the stimulus but also a prospective action plan. Such entangled tuning for different cognitive variables has been observed repeatedly in the PFC [120] and has been proposed as a general model of PFC function [121].

The transformation of sensory signals into an intermediate format that bridges between stimuli and delayed target responses can be easily appreciated. Consider a task that involves memorizing a picture of a lion and selecting this animal from one of two alternative targets after a delay. The incoming target stimulus to which a memorized sample must be matched will drive the

---

Box 2. A Hierarchy of Spatial Scale

Spatial maps can be found in visual, parietal, and frontal cortices alike [158–160]. This reflects the predominance of space as an organizational principle across many brain areas. The encoding of memorized positions has also been observed across many stages of the cortical hierarchy including in the lPFC [31,61,62,76,87]. As for nonspatial features, there is a gradient of representation across the hierarchy: (i) higher-level areas exhibit larger spatial receptive fields [161–163]; (ii) spatial mnemonic representations in parietal and frontal cortices are closely tied to the preparation of actions and the control of attention shifts [164–166]; and (iii) spatial representations in higher-level regions generalize across many different sensory modalities [167–169] and involve nonretinal frames of reference [135,170]. In contrast, lower-level maps are recruited when tasks demand memory of fine spatial detail in retinal coordinates [50].

Mnemonic maps of visual space could also be involved in the representation of other feature information. For example, real-life objects are not spatially homogeneous and are known to elicit differential distributions of attention and saliency across space, even for objects of similar size [171]. The frontal eye fields and parietal cortex encode the spatial distribution of saliency within natural images [172,173]. Thus, information that appears to be specific to the nonspatial features of an object might also depend on spatial memory of particularly salient subregions within the shape of the object.

Stimulus-selective information in the PFC has been shown for low-level nonspatial features such as grating orientation [77] and motion direction [29,49]. Recent behavioral and electrophysiological work suggests that spatial strategies might play a role in memorizing such features [174]. Consistent with this possibility, orientation-selective responses in V1 correlate with the topography of spatial selectivity, demonstrating joint selectivity for orientation and space [175]. In line with this, selectivity for orientation, motion, and space during working memory largely overlap in parietal and frontal cortices (see Figure 1). Hence, feature-selective representations during working memory could arise at multiple representational scales of space that are distributed across the cortical hierarchy.

---

sensory cortex in a similar way as the sample. Thus, their traces would likely interfere if they were maintained in the same low-level area (see [43,94–97]). A solution would be to transform the representation and store it in a manner that allows the comparison to subsequent stimuli. For example, the relevant features of the lion could be maintained in the PFC as a feature-selective attentional search template [122,123]. This would have the additional advantage of avoiding interference between sample and target stimuli, avoiding that they overwrite the representation of the sample [124]. If this is the case, the nature of stimulus representations in PFC should be strongly task dependent [123]. Finding a yellow lion among yellow giraffes, for instance, benefits from an attentional focus on shape rather than color, whereas finding a yellow lion among grey elephants will benefit more from an attentional focus on color rather than shape ([125], Figure 2C).

The transformation of a sensory item into a behavioral response can be observed at multiple hierarchical levels of the brain, ranging from sensory to prefrontal cortices [122,123,126–128], and culminates in concrete preparatory motor activity in the premotor and primary motor cortices [129,130]. Furthermore, as this process unfolds during the delay-period, task-related preparation signals are expected to gradually increase in strength and be strongest immediately before the onset of the test stimulus. This is indeed what is found in the PFC; at least in a subset of cells [6,43] (see also [95,131]), which suggests that these forms of persistent activity can also play a preparatory role, going beyond the memorization of sensory information. Such a response-oriented, prospective nature of representations in the lateral prefrontal cortex (lPFC) is further supported by delayed paired association tasks [132], where animals are required to learn an association between a sample and test stimulus, separated by a brief delay. Under these conditions, PFC neurons encode the transformation from a memory representation of the sample stimulus to a representation of the expected paired associate in preparation of the test stimulus presentation.

A Working Memory Gradient

We conceptualize working memory as a hierarchical process that links detailed sensory representations to specific behavioral responses via intermediate task-relevant representations and action plans in a network of brain areas that each represent the working memory in a format

that matches their functional specialization (see also [15]). Thus, we propose that every cortical region can produce persistent stimulus-selective activity if the features that are coded in that region need to be memorized across a delay to prepare for an upcoming response. Such a universal cortical capability to produce persistent activity explains a simple observation that may not have received the attention it deserves, and which we here put forward as a central hypothesis: all features that are represented in the neocortex can be briefly memorized. Violations of this rule would provide important insights into the mechanisms underlying working memory. As far as we know, however, this proposal holds for low-level as well as high-level sensory features, which can be stored as persistent activity in lower and higher areas of the cerebral cortex, respectively. It also holds for transformed representations that are used to guide upcoming behavior.

The two potential explanations for the distributed nature of working memory outlined above are not mutually exclusive. In fact, they both roughly map on a similar posterior-to-frontal axis of functional brain organization. At the posterior end, the sensory cortices represent incoming sensory information in a relatively pure and detailed form. At the frontal end of the gradient, the frontal cortex represents information that is abstracted and transformed in support of upcoming behavior. Some of the frontal representations are abstract and categorical [109], however, other representations more directly support the execution of actions [15,133], and code for remembered action features such as the location of objects in eye coordinates for eye movements [134], in arm coordinates for hand movements [129], or even in object coordinates if this is required for the task [135].

Areas in between the sensory and prefrontal cortices might maintain intermediate transformations in many different formats. The gradients mentioned can be observed at multiple stages and reflect the increasing levels of processing and abstraction across subsequent stages of stimulus processing. One major visual stream, for example, originates from the primary visual cortex, and then proceeds via the extrastriate and inferior temporal cortices to the lPFC. It is well established that the tuning of single cells in these regions exhibits increasing degrees of spatial invariance, stimulus generalization, and abstraction. Whereas cells in V1 are selective to the detailed low-level features of stimuli, cells in inferior temporal cortex exhibit translational invariance and they can generalize across different low-level features [136]. Consistently, inferior temporal cells exhibit persistent stimulus-selective activity for objects during working memory delays [54]. Our review of the literature (see Figure 1) is generally consistent with the notion of a gradient with increasing levels of abstraction of memory signals across different levels of processing.

Which nodes maintain working memory representations depends on which format is the most suitable to perform the task at hand. When, for instance, detailed sensory information is required to perform a task, it makes sense to store stimulus information in areas suitable to represent such details, like the early sensory cortex. In contrast, when a task promotes the use of categorical information, it may be stored in a more abstract or verbal format by neurons higher in the processing hierarchy. Finally, if the appropriate response is known as soon as a sample stimulus is shown, working memory can also be encoded in the form of a pure response plan [137].

The notion of a distributed working memory network that postulates the ability of retention as a general principle in the neocortex has a long history [133,138]. Already early on, behavioral studies have revealed the effects of varying levels of processing in memory [139] and have distinguished between stores for different stimulus modalities [1]. More recent accounts of distributed working memory have emphasized different timescales of retention in different brain areas [140,141], the role of attentional processes in the retention of memories [4,11,142], and the utility of synaptic plasticity and dynamic firing trajectories for working memory [2,6,32,35,143].

Our hypothesis is ignorant about the precise mechanism for persistent activity: it might be generated within the cortex itself or rely on recurrent loops through subcortical structures, such as the thalamus or basal ganglia [144–147]. Furthermore, the working memory gradient outlined above does not require a separate homunculus control mechanism [145,146,148], but can instead be produced by the underlying neural circuits themselves, similar to the idea of a **Turing machine** ([149], Box 3). Indeed, recent computational studies have reported biologically plausible learning rules that can induce persistent activity for working memory, specifically for those features that need to be kept online during memory delays [150,151].

The ability of the brain to simultaneously encode working memory representations in multiple regions and at different levels of abstraction and response preparation necessarily involves redundancy. One advantage of redundancy is that it strongly increases the robustness of working memory representations. One study investigated simultaneous representations of a motor plan in the left and right premotor cortex. It was found that memory signals could be restored from the contralateral (i.e., redundant) side when one side was optogenetically silenced [152]. More generally, simultaneous and partially redundant representations might reflect the effects of representational loops that have been postulated for visual (attention-based rehearsal, [11]) and verbal (phonological loop, [1]) working memory.

Given the abundance of neurons capable of persistent activity, it is remarkable that the capacity to retain information in working memory is severely limited [153,154]. Hence, there must exist restrictions on the number of coactive memories per node and/or the number of coactive nodes. While individual stores might already have capacity restrictions [46,155], we suggest that properties of the overall working memory network might contribute to limitations of capacity as well (see Box 4). Specifically, if the maintenance of every individual memory item relies on the

---

Box 3. Distributed Control of the Working Memory Network

Here, we propose that signals in all cortical regions can exhibit a short-term buffering of information, depending on the nature of the task. Accordingly, working memory contents can range from low-level to abstract features and from sensory to motor codes. This raises the question how the distributed storage across multiple regions is coordinated. Theoretically, it might be conceivable that one or a few regions specifically control encoding, storage and retention, similar to a central executive [1]. However, it is also conceivable that control itself is distributed [176,177]. On the one hand, prefrontal regions can exert top-down attentional control and influence which low-level features are selected to be encoded in working memory. On the other hand, a low-level sensory cue (or its working memory) encoded in the sensory cortex can define a task and thus influence activity in the PFC. Thus, control cannot only go top-down but also bottom-up, and some sense of control thus seems to be present across the entire hierarchy.

The proposal that every brain region can exert an influence over processing in (most) other regions when it is required by the task is conceptually related to models that only involve states and transitions between states. Such models, for instance Markov processes [178] or Turing machines [179], do not require a centralized control system. A Turing machine is a hypothetical apparatus developed by Alan Turing to address the nature of computational operations. Turing machines consist of a finite-state machine and an infinite tape. Actions are selected based on the internal state of the machine and the symbol of the tape currently 'seen' by the machine. Selected actions then cause the machine to change its internal state by writing a new symbol to the tape, and/or shifting the tape to reveal the next symbol. The computational versatility of Turing machines arises from this continuous cycle of actions and the updating of internal states and 'tape memory'. Importantly, the tape can specify simple input but also entire programs using a similar format. Clearly, there are many differences between Turing machines and brains (e.g., Turing machines are serial and lack the distinction between long-term and working memory). Yet, there are also interesting analogies with working memory processes in the brain [149]. For instance, working memories are internal neural states, which may encode previous sensory inputs or even complex task rules. These memories in turn influence the selected actions. Motor actions influence incoming sensory information (e.g., arm movements influence proprioception and eye movements change visual input). Likewise, internal actions, like attention shifts or the recall of associations, modulate active mnemonic codes in the frontal cortex and lower brain areas. Thus, selected actions can impact the pattern of sensory input, the activated mnemonic codes, or both (just as actions of Turing machines change the internal state, the tape, or both). Consequentially, new actions can be selected at the next time step, causing new changes in sensory input and working memory contents. This succession of action selection and updating of sensory and mnemonic codes in working memory resembles the workings of a Turing machine.

Box 4. The Limited Capacity of Working Memory

The wide distribution of persistent stimulus-selective responses in the primate brain yields the question whether capacity limitations in working memory reflect properties of individual storage regions, or a property of the distributed network. Behavioral models assume either a slot-based [180,181] or resource-based [182,183] limitation of capacity. At a neural level, several factors might limit memory capacity.

Competition for representation: multiple items may compete for representation within the same representational maps (e.g., of visual space in V1) or in any other limited 'cortical real estate' [155], such that individual memorized items interfere with each other during the delay period or already during encoding [46]. In line with this, individual visual working memory capacity is correlated with the volumetric size of V1 [184], and persistent stimulus-selective activity per memory item decreases with increasing load [12]. Capacity increases with the sparseness of neural representations that leads to less interference between representations [40]. It can be further improved by synaptic facilitation [40]. In contrast, capacity might be a global rather than local property of the working memory network. For example, memorization of low-level visual features reduces the ability to retain low-level auditory information [185].

Distributed storage: the above-mentioned interference among memory representations could decrease if different individual stimuli of the same type (e.g., multiple oriented gratings) were represented in different brain areas. Memorized stimuli that are currently relevant for an ongoing task and that are in the focus of attention [186] might, for instance, be retained by detailed representations in sensory cortex, while secondary memory items that are stored for later use are retained in a coarser, more abstract form in anterior regions [187]. This would explain why persistent stimulus-selective activity in sensory regions is weaker or absent for items that are stored outside the focus of attention, while it is reinstated as soon as they regain their relevance [30,76,188].

Task-relevant detail: the level of detail required for a working memory task has implications for capacity. The posterior-to-frontal gradient of increasing abstraction proposed here (see 'Different Levels of Abstraction') can be interpreted as a form of complexity reduction [189] or 'chunking' [153]. For example, 'd', 'o', and 'g' are more easily retained as a word ('dog') than as individual letters [190]. Similarly, and in accordance with the flexible allocation of memory resources [182,183,191], higher cortical areas can retain larger chunks of low-level features that would be difficult to store individually in lower level regions. However, such chunking involves a step of abstraction that might discard low-level visual features.

---

activity within a widespread network, the representations of multiple memoranda will more readily interfere with each other. Consequently, the selection of the nodes that will retain stimulus- or response-selective information across a delay interval should strongly depend on task requirements [80,88] to avoid interference (e.g., from distractors, [43,94–97]) and thereby boost memory capacity (see Outstanding Questions).

Concluding Remarks

There is abundant evidence for widely distributed stimulus-related information in sensory, parietal, and prefrontal cortices during working memory delays. In the absence of evidence implicating any of these regions as an exclusive and localized store of memory contents, we suggest that working memory is better characterized as a distributed network that gradually transforms sensory information towards an appropriate behavioral response, across a temporal delay. Persistent stimulus-specific activity might be observed anywhere in such a network and at any stage of transformation. Localization will thus strongly depend on the precise requirements and context of the task. This notion suggests that perhaps the field of working memory should shift its focus from asking where in the brain working memories are stored to unraveling how a range of highly specialized brain areas together transform a sensory stimulus into an appropriate response and how this process is sustained as a working memory across delays.

Outstanding Questions

What is the mechanism of retention? Two main neurocomputational mechanisms of retention have been proposed: (i) activity-based retention, such as spiking activity, reverberatory activity, feed-forward-connected subgroups of neurons (synfire chains), oscillatory dynamics, or high-dimensional dynamic trajectories; and (ii) activity-silent retention, where stimulus-selective representations are maintained as a pattern of synaptic weights. These two mechanisms are not mutually exclusive and might jointly contribute to retention.

How does the current relevance of a memorized item affect its retention? Behavioral studies suggest a distinction between attended and unattended (accessory) items held in memory, but persistent stimulus-specific activity has only been reported for attended items. Unattended representations might be retained (i) as activity-silent synaptic representations, (ii) using a different neural code than attended representations (e.g., on different levels of abstraction; see Box 4), or (iii) as weaker activity patterns then attended ones.

How can we distinguish between true mnemonic activity of a feature and encoding of other stimulus-related factors? Different sensory stimuli are known to be associated with differences in valence, motor affordance, and ease of recognition. All these features are known to affect signals across the neocortex. Thus, signals that appear to be stimulus selective might instead reflect encoding of these other features.

How is the retention of multiple items coordinated and how do these memories interact? Storing stimulus information in memory in multiple formats and areas in parallel can increase working memory robustness. However, to date it is unclear: (i) how multiple memory representations interact, (ii) which mechanism selects which memory will guide behavior, (iii) how parallel representations can compensate for perturbations and interference (e.g., by distracting stimuli), and (iv) how the retention of a particular working memory content is terminated.

Supplemental Information

Supplemental information associated with this article can be found online at http://dx.doi.org/10.1016/j.tics.2016.12.007.

References

1. Baddeley, A.D. (1986) *Working Memory,* Clarendon Press
2. D'Esposito, M. and Postle, B.R. (2015) The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* 66, 115–142
3. Goldman-Rakic, P. (1995) Cellular basis of working memory. *Neuron* 14, 477–485
4. Postle, B.R. (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38
5. Sreenivasan, K.K. *et al.* (2014) Revisiting the role of persistent neural activity during working memory. *Trends Cogn. Sci.* 18, 82–89
6. Stokes, M.G. (2015) "Activity-silent" working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* 19, 394–405
7. Xu, Y. and Jeong, S.K. (2015) The contribution of human superior intraparietal sulcus to visual short-term memory and perception. In *Mechanisms of Sensory Working Memory: Attention and Perfomance XXV*. Elsevier
8. Riley, M.R. and Constantinidis, C. (2016) Role of prefrontal persistent activity in working memory. *Front. Syst. Neurosci.* 9, 181
9. Durstewitz, D. *et al.* (2000) Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* 83, 1733–1750
10. Mendoza-Halliday, D. *et al.* (2015) Working memory representations of visual motion along the primate dorsal visual pathway. In *Mechanisms of Sensory Working Memory: Attention and Perfomance XXV*. Elsevier
11. Awh, E. and Jonides, J. (2001) Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.* 5, 119–126
12. Emrich, S.M. (2015) What are the roles of sensory and parietal activity in visual short-term memory? In *Mechanisms of Sensory Working Memory: Attention and Perfomance XXV*. Elsevier
13. Pasternak, T. *et al.* (2015) Unilateral prefrontal lesions impair memory-guided comparisons of contralateral visual motion. *J. Neurosci.* 35, 7095–7105
14. Fuster, J.M. (1997) Network memory. *Trends Neurosci.* 20, 451–459
15. Zimmer, H.D. (2008) Visual and spatial working memory: from boxes to networks. *Neurosci. Biobehav. Rev.* 32, 1373–1395
16. Eriksson, J. *et al.* (2015) Neurocognitive architecture of working memory. *Neuron* 88, 33–46
17. Lee, S.-H. and Baker, C.I. (2016) Multi-voxel decoding and the topography of maintained information during visual working memory. *Front. Syst. Neurosci.* 10, 2
18. Jacobsen, C.F. (1936) The functions of the frontal association areas in monkeys. *Comparative Psychology Monographs* 13, 1–60
19. Warrington, E.K. and Shallice, T. (1969) The selective impairment of auditory verbal short-term memory. *Brain* 92, 885–896
20. D'Esposito, M. and Postle, B.R. (1999) The dependence of span and delayed-response performance on prefrontal cortex. *Neuropsychologia* 37, 1303–1315
21. Malmo, R.B. (1942) Interference factors in delayed response in monkeys after removal of frontal lobes. *J. Neurophysiol.* 5, 295–308
22. Mishkin, M. (1957) Effects of small frontal lesions on delayed alternation in monkeys. *J. Neurophysiol.* 20, 615–622
23. Goldman, P.S. and Rosvold, H.E. (1970) Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. *Exp. Neurol.* 27, 291–304
24. Petrides, M. (1995) Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *J. Neurosci.* 15, 359–375
25. Levy, R. and Goldman-Rakic, P.S. (1999) Association of storage and processing functions in the dorsolateral prefrontal cortex of the nonhuman primate. *J. Neurosci.* 19, 5149–5158
26. Fuster, J.M. and Alexander, G.E. (1971) Neuron activity related to short-term memory. *Science* 173, 652–654
27. Wager, T.D. and Smith, E.E. (2003) Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274
28. Pasternak, T. and Greenlee, M.W. (2005) Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6, 97–107
29. Zaksas, D. and Pasternak, T. (2006) Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *J. Neurosci.* 26, 11726–11742
30. Lewis-Peacock, J.A. *et al.* (2012) Neural evidence for a distinction between short-term memory and the focus of attention. *J. Cogn. Neurosci.* 24, 61–79
31. Lara, A.H. and Wallis, J.D. (2014) Executive control processes underlying multi-item working memory. *Nat. Neurosci.* 17, 876–883
32. Mongillo, G. *et al.* (2008) Synaptic theory of working memory. *Science* 319, 1543–1546
33. Sugase-Miyamoto, Y. *et al.* (2008) Short-term memory trace in rapidly adapting synapses of inferior temporal cortex. *PLoS Comput. Biol.* 4, e1000073
34. Barak, O. and Tsodyks, M. (2014) Working models of working memory. *Curr. Opin. Neurobiol.* 25, 20–24
35. Stokes, M.G. *et al.* (2013) Dynamic coding for cognitive control in prefrontal cortex. *Neuron* 78, 364–375
36. Sreenivasan, K.K. *et al.* (2014) Distributed and dynamic storage of working memory stimulus information in extrastriate cortex. *J. Cogn. Neurosci.* 26, 1141–1153
37. Lundqvist, M. *et al.* (2016) Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164
38. Hebb, D.O. (1949) *The Organization of Behavior: A Neuropsychological Theory,* Wiley
39. Amit, D.J. (1995) The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behav. Brain Sci.* 18, 617–626
40. Rolls, E.T. *et al.* (2013) Holding multiple items in short term memory: a neural mechanism. *PLoS One* 8, e61078
41. Kubota, K. *et al.* (1974) Visuokinetic activities of primate prefrontal neurons during delayed-response performance. *J. Neurophysiol.* 37, 1197–1212
42. Rainer, G. *et al.* (1998) Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* 393, 577–578
43. Miller, E.K. *et al.* (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* 16, 5154–5167
44. Meyer, T. *et al.* (2011) Stimulus selectivity in dorsal and ventral prefrontal cortex after training in working memory tasks. *J. Neurosci.* 31, 6266–6276
45. Rao, S.C. *et al.* (1997) Integration of what and where in the primate prefrontal cortex. *Science* 276, 821–824
46. Buschman, T.J. *et al.* (2011) Neural substrates of cognitive capacity limitations. *Proc. Natl. Acad. Sci. U. S. A.* 108, 11252–11255

47. Nieder, A. *et al.* (2002) Representation of the quantity of visual items in the primate prefrontal cortex. *Science* 297, 1708–1711

48. Romo, R. *et al.* (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399, 470–473

49. Mendoza-Halliday, D. *et al.* (2014) Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat. Neurosci.* 17, 1255–1262

50. van Kerkoerle, T. *et al.* The influence of attention and working memory on neuronal activity in the different layers of primary visual cortex. *Nat. Commun.* (in press).

51. Supèr, H. *et al.* (2001) A neural correlate of working memory in the monkey primary visual cortex. *Science* 293, 120–124

52. Bisley, J.W. *et al.* (2004) Activity of neurons in cortical area MT during a memory for motion task. *J. Neurophysiol.* 91, 286–300

53. Miyashita, Y. and Chang, H.S. (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* 331, 68–70

54. Miller, E.K. *et al.* (1993) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J. Neurosci.* 13, 1460–1478

55. Chelazzi, L. *et al.* (2001) Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb. Cortex* 11, 761–772

56. Fuster, J.M. and Jervey, J.P. (1982) Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *J. Neurosci.* 2, 361–375

57. Fuster, J. and Jervey, J. (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952–955

58. Gottlieb, Y. *et al.* (1989) Single unit activity in the auditory cortex of a monkey performing a short term memory task. *Exp. Brain Res.* 74, 139–148

59. Zhou, Y.D. and Fuster, J.M. (1996) Mnemonic neuronal activity in somatosensory cortex. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10533–10537

60. Hernández, A. *et al.* (2010) Decoding a perceptual decision process across cortex. *Neuron* 66, 300–314

61. Armstrong, K.M. *et al.* (2009) Selection and maintenance of spatial information by frontal eye field neurons. *J. Neurosci.* 29, 15621–15629

62. Constantinidis, C. and Steinmetz, M.A. (1996) Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task. *J. Neurophysiol.* 76, 1352–1355

63. Vergara, J. *et al.* (2016) A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron* 89, 54–62

64. Suzuki, W.A. *et al.* (1997) Object and place memory in the macaque entorhinal cortex. *J. Neurophysiol.* 78, 1062–1081

65. Watanabe, Y. and Funahashi, S. (2012) Thalamic mediodorsal nucleus and working memory. *Neurosci. Biobehav. Rev.* 36, 134–142

66. Horwitz, G.D. and Newsome, W.T. (1999) Separate signals for target selection and movement specification in the superior colliculus. *Science* 284, 1158–1161

67. Alexander, G.E. (1987) Selective neuronal discharge in monkey putamen reflects intended direction of planned limb movements. *Exp. Brain Res.* 67, 623–634

68. Prut, Y. and Fetz, E.E. (1999) Primate spinal interneurons show pre-movement instructed delay activity. *Nature* 401, 590–594

69. Courtney, S.M. *et al.* (1996) Object and spatial visual working memory activate separate neural systems in human cortex. *Cereb. Cortex* 6, 39–49

70. Ranganath, C. *et al.* (2004) Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J. Neurosci.* 24, 3917–3925

71. Haxby, J.V. *et al.* (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430

72. Haynes, J.-D. and Rees, G. (2006) Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534

73. Kamitani, Y. and Tong, F. (2005) Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–685

74. Kriegeskorte, N. *et al.* (2006) Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3863–3868

75. Brouwer, G.J. and Heeger, D.J. (2009) decoding and reconstructing color from responses in human visual cortex. *J. Neurosci.* 29, 13992–14003

76. Sprague, T.C. *et al.* (2014) Reconstructions of information in visual spatial working memory degrade with memory load. *Curr. Biol.* 24, 2174–2180

77. Ester, E.F. *et al.* (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* 87, 893–905

78. Ester, E.F. *et al.* (2013) A neural measure of precision in visual working memory. *J. Cogn. Neurosci.* 25, 754–761

79. Harrison, S.A. and Tong, F. (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635

80. Serences, J.T. *et al.* (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* 20, 207–214

81. Christophel, T.B. *et al.* (2012) Decoding the contents of visual short-term memory from human visual and parietal cortex. *J. Neurosci.* 32, 12983–12989

82. Riggall, A.C. and Postle, B.R. (2012) The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J. Neurosci.* 32, 12990–12998

83. Pratte, M.S. and Tong, F. (2014) Spatial specificity of working memory representations in the early visual cortex. *J. Vis.* 14, 22

84. Linke, A.C. *et al.* (2011) Stimulus-specific suppression preserves information in auditory short-term memory. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12961–12966

85. Kumar, S. *et al.* (2016) A brain system for auditory working memory. *J. Neurosci.* 36, 4492–4505

86. Christophel, T.B. and Haynes, J.-D. (2014) Decoding complex flow-field patterns in visual working memory. *Neuroimage* 91, 43–51

87. Jerde, T.A. *et al.* (2012) Prioritized maps of space in human frontoparietal cortex. *J. Neurosci.* 32, 17382–17390

88. Lee, S.-H. *et al.* (2013) Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat. Neurosci.* 16, 997–999

89. Spitzer, B. *et al.* (2010) Oscillatory correlates of vibrotactile frequency processing in human working memory. *J. Neurosci.* 30, 4496–4502

90. Spitzer, B. and Blankenburg, F. (2011) Stimulus-dependent EEG activity reflects internal updating of tactile working memory in humans. *Proc. Natl. Acad. Sci. U. S. A.* 108, 8444–8449

91. Spitzer, B. and Blankenburg, F. (2012) Supramodal parametric working memory processing in humans. *J. Neurosci.* 32, 3287–3295

92. Christophel, T. *et al.* (2013) Decoding invariant representations in visual working memory. *J. Vis.* 13, 927

93. Yan, C. *et al.* (2015) Decoding neural representations of Chinese characters stored in working memory. In *Presented at the Annual Meeting of the Organization for Human Brain Mapping*. Organization for Human Brain Mapping, Honululu

94. Woloszyn, L. and Sheinberg, D.L. (2009) Neural dynamics in inferior temporal cortex during a visual working memory task. *J. Neurosci.* 29, 5494–5507

95. Hayden, B.Y. and Gallant, J.L. (2013) Working memory and decision processes in visual area V4. *Front. Neurosci.* 7, 18

96. Jacob, S.N. and Nieder, A. (2014) Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. *Neuron* 83, 226–237

97. Bettencourt, K.C. and Xu, Y. (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat. Neurosci.* 19, 150–157

98. Glanzer, M. and Clark, W.H. (1964) The verbal-loop hypothesis: conventional figures. *Am. J. Psychol.* 77, 621–626

99. Haber, R.N. (1964) Effects of coding strategy on perceptual memory. *J. Exp. Psychol.* 68, 357–362

100. Murdock, B.B. (1967) Recent developments in short-term memory. *Brit. J. Psychol.* 58, 421–433

101. Murray, D.J. (1965) The effect of white noise upon the recall of vocalized lists. *Can. J. Psychol.* 19, 333–345

102. Waugh, N.C. and Norman, D.A. (1965) Primary memory. *Psychol. Rev.* 72, 89–104

103. Cermak, G.W. (1971) Short-term recognition memory for complex free-form figures. *Psychon. Sci.* 25, 209–211

104. Posner, M.I. and Keele, S.W. (1967) Decay of visual information from a single letter. *Science* 158, 137–139

105. Coltheart, V. (1972) The effects of acoustic and semantic similarity on concept identification. *Q. J. Exp. Psychol.* 24, 55–65

106. Logie, R.H. *et al.* (2000) Visual similarity effects in immediate verbal serial recall. *Q. J. Exp. Psychol. A* 53, 626–646

107. Miller, E.K. *et al.* (2003) Neural correlates of categories and concepts. *Curr. Opin. Neurobiol.* 13, 198–203

108. Bae, G.-Y. *et al.* (2015) Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *J. Exp. Psychol. Gen.* 144, 744

109. Freedman, D.J. *et al.* (2001) Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291, 312–316

110. Freedman, D.J. *et al.* (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J. Neurosci.* 23, 5235–5246

111. Spitzer, B. *et al.* (2014) Working memory coding of analog stimulus properties in the human prefrontal cortex. *Cereb. Cortex* 24, 2229–2236

112. Romo, R. and de Lafuente, V. (2013) Conversion of sensory signals into perceptual decisions. *Prog. Neurobiol.* 103, 41–75

113. Nieder, A. (2016) The neuronal code for number. *Nat. Rev. Neurosci.* 17, 366–382

114. Nieder, A. (2012) Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. U. S. A.* 109, 11860–11865

115. Spitzer, B. *et al.* (2014) Parametric alpha- and beta-band signatures of supramodal numerosity information in human working memory. *J. Neurosci.* 34, 4293–4302

116. Lara, A.H. and Wallis, J.D. (2015) The role of prefrontal cortex in working memory: a mini review. *Front. Syst. Neurosci.* 9, 173

117. Bancroft, T.D. *et al.* (2014) Does stimulus complexity determine whether working memory storage relies on prefrontal or sensory cortex? *Atten. Percept. Psychophys.* 76, 1954–1961

118. Ku, Y. *et al.* (2015) Prefrontal cortex and sensory cortices during working memory: quantity and quality. *Neurosci. Bull.* 31, 175–182

119. Mansouri, F.A. *et al.* (2015) Working memory in the service of executive control functions. *Front. Syst. Neurosci.* 9, 166

120. Asaad, W.F. *et al.* (1998) Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21, 1399–1407

121. Rigotti, M. *et al.* (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497, 585–590

122. Zhou, H. and Desimone, R. (2011) Feature-based attention in the frontal eye field and area V4 during visual search. *Neuron* 70, 1205–1217

123. Bichot, N.P. *et al.* (2015) A source for feature-based attention in the prefrontal cortex. *Neuron* 88, 832–844

124. Miller, E.K. and Desimone, R. (1994) Parallel neuronal mechanisms for short-term memory. *Science* 263, 520–522

125. Squire, R.F. *et al.* (2013) Prefrontal contributions to visual selective attention. *Annu. Rev. Neurosci.* 36, 451–466

126. Chelazzi, L. *et al.* (1993) A neural basis for visual search in inferior temporal cortex. *Nature* 363, 345–347

127. Luck, S.J. *et al.* (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* 77, 24–42

128. Peelen, M.V. and Kastner, S. (2011) A neural basis for real-world visual search in human occipitotemporal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12125–12130

129. Graziano, M.S. *et al.* (1994) Coding of visual space by premotor neurons. *Science* 266, 1054–1057

130. Gallivan, J.P. *et al.* (2011) Decoding effector-dependent and effector-independent movement intentions from human parieto-frontal brain activity. *J. Neurosci.* 31, 17149–17168

131. Qi, X.-L. *et al.* (2010) Comparison of neural activity related to working memory in primate dorsolateral prefrontal and posterior parietal cortex. *Front. Syst. Neurosci.* 4, 12

132. Rainer, G. *et al.* (1999) Prospective coding for objects in primate prefrontal cortex. *J. Neurosci.* 19, 5493–5505

133. Fuster, J.M. (1995) *Memory in the Cerebral Cortex: An Empirical Approach to Neural Networks in the Human and Nonhuman Primate,* MIT Press

134. Funahashi, S. *et al.* (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349

135. Olson, C.R. and Gettner, S.N. (1995) Object-centered direction selectivity in the macaque supplementary eye field. *Science* 269, 985–988

136. Sary, G. *et al.* (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260, 995–997

137. Constantinidis, C. *et al.* (2001) The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nat. Neurosci.* 4, 311–316

138. Ghent, L. *et al.* (1962) Short-term memory after frontal-lobe injury in man. *J. Comp. Physiol. Psychol.* 55, 705

139. Craik, F.I.M. and Lockhart, R.S. (1972) Levels of processing: a framework for memory research. *J. Verb. Learn. Verb. Be.* 11, 671–684

140. Murray, J.D. *et al.* (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* 17, 1661–1663

141. Hasson, U. *et al.* (2015) Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn. Sci.* 19, 304–313

142. Jonides, J. *et al.* (2008) The mind and brain of short-term memory. *Annu. Rev. Psychol.* 59, 193–224

143. Serences, J.T. (2016) Neural mechanisms of information storage in visual short-term memory. *Vis. Res.* 128, 53–67

144. Wang, X.-J. (2001) Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24, 455–463

145. Hazy, T.E. *et al.* (2006) Banishing the homunculus: making working memory work. *Neuroscience* 139, 105–118

146. Hazy, T.E. *et al.* (2007) Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. *Phil. Trans. R. Soc. B* 362, 1601–1613

147. Gazzaley, A. and Nobre, A.C. (2012) Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.* 16, 129–135

148. Baddeley, A. (2012) Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29

149. Zylberberg, A. *et al.* (2011) The human Turing machine: a neural framework for mental programs. *Trends Cogn. Sci.* 15, 293–300

150. Rombouts, J.O. *et al.* (2015) How attention can create synaptic tags for the learning of working memories in sequential tasks. *PLoS Comput. Biol.* 11, e1004060

151. O'Reilly, R.C. and Frank, M.J. (2006) Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* 18, 283–328

152. Li, N. *et al.* (2016) Robust neuronal dynamics in premotor cortex during motor planning. *Nature* 532, 459–464

153. Miller, G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 101, 343–352

154. Cowan, N. (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114

155. Franconeri, S.L. *et al.* (2013) Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cogn. Sci.* 17, 134–141

156. Müller, N.G. and Knight, R.T. (2006) The functional neuroanatomy of working memory: contributions of human brain lesion studies. *Neuroscience* 139, 51–58

157. Britten, K.H. *et al.* (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* 13, 87–100

158. Lamme, V.A. and Roelfsema, P.R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579

159. Hagler, D.J., Jr and Sereno, M.I. (2006) Spatial maps in frontal and prefrontal cortex. *Neuroimage* 29, 567–577

160. Wandell, B.A. *et al.* (2007) Visual field maps in human cortex. *Neuron* 56, 366–383

161. Dumoulin, S.O. and Wandell, B.A. (2008) Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647–660

162. Sprague, T.C. and Serences, J.T. (2013) Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nat. Neurosci.* 16, 1879–1887

163. Klein, B.P. *et al.* (2014) Attraction of position preference by spatial attention throughout human visual cortex. *Neuron* 84, 227–237

164. Tootell, R.B. *et al.* (1998) The retinotopy of visual spatial attention. *Neuron* 21, 1409–1422

165. Culham, J.C. *et al.* (2006) The role of parietal cortex in visuomotor control: what have we learned from neuroimaging? *Neuropsychologia* 44, 2668–2684

166. Shi, Y. *et al.* (2013) Multimodal representation of limb endpoint position in the posterior parietal cortex. *J. Neurophysiol.* 109, 2097–2107

167. Andersen, R.A. *et al.* (1997) Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu. Rev. Neurosci.* 20, 303–330

168. Schlack, A. *et al.* (2005) Multisensory space representations in the macaque ventral intraparietal area. *J. Neurosci.* 25, 4616–4625

169. Breveglieri, R. *et al.* (2008) Visual, somatosensory, and bimodal activities in the macaque parietal area PEc. *Cereb. Cortex* 18, 806–816

170. Graziano, M.S.A. *et al.* (1997) Coding the locations of objects in the dark. *Science* 277, 239–241

171. Yun, K. *et al.* (2013) Exploring the role of gaze behavior and object detection in scene understanding. *Front. Psychol.* 4, 917

172. Gottlieb, J.P. *et al.* (1998) The representation of visual salience in monkey parietal cortex. *Nature* 391, 481–484

173. Bogler, C. *et al.* (2011) Decoding successive computational stages of saliency processing. *Curr. Biol.* 21, 1667–1671

174. Fahrenfort, J. *et al.* (2016) Alpha-band and raw EEG reflect distinct maintenance mechanisms during working memory. *J. Vis.* 16, 711

175. Freeman, J. *et al.* (2011) Orientation decoding depends on maps, not columns. *J. Neurosci.* 31, 4792–4804

176. Collette, F. and Van der Linden, M. (2002) Brain imaging of the central executive component of working memory. *Neurosci. Biobehav. Rev.* 26, 105–125

177. Monsell, S. and Driver, J. (2000) Banishing the control homunculus. In *In Control of Cognitive Processes: Attention and Performance XVIII*, pp. 3–32, MIT Press

178. Norris, J.R. (1997) *Markov Chains,* Cambridge University Press

179. Turing, A.M. (1937) On computable numbers, with an application to the entscheidungs problem. *Proc. London Math. Soc.* s2-42, 230–265

180. Luck, S.J. and Vogel, E.K. (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–280

181. Luck, S.J. and Vogel, E.K. (2013) Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400

182. Bays, P.M. *et al.* (2009) The precision of visual working memory is set by allocation of a shared resource. *J. Vis.* 9, 1–11

183. Ma, W.J. *et al.* (2014) Changing concepts of working memory. *Nat. Neurosci.* 17, 347–356

184. Bergmann, J. *et al.* (2016) Neural anatomy of primary visual cortex limits visual working memory. *Cereb. Cortex* 26, 43–50

185. Salmela, V.R. *et al.* (2014) Working memory resources are shared across sensory modalities. *Atten. Percept. Psychophys.* 76, 1–13

186. Cowan, N. (1995) *Attention and Memory: An Integrated Framework,* Oxford University Press

187. Olivers, C.N.L. *et al.* (2011) Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* 15, 327–334

188. LaRocque, J.J. *et al.* (2012) Decoding attended information in short-term memory: an EEG study. *J. Cogn. Neurosci.* 25, 127–142

189. Riesenhuber, M. and Poggio, T. (1999) Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025

190. Simon, H.A. (1974) How big is a chunk? By combining data from several experiments, a basic human memory unit can be identified and measured. *Science* 183, 482–488

191. Alvarez, G.A. and Cavanagh, P. (2004) The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol. Sci.* 15, 106–111

192. Albers, A.M. *et al.* (2013) Shared representations for working memory and mental imagery in early visual cortex. *Curr. Biol.* 23, 1427–1431

193. Christophel, T.B. *et al.* (2015) Parietal and early visual cortices encode working memory content across mental transformations. *Neuroimage* 106, 198–206

194. Wimmer, K. *et al.* (2016) Prefrontal neurons represent motion signals from across the visual field but for memory-guided comparisons depend on neurons providing these signals. *J. Neurosci.* 36, 9351–9364

195. Freedman, D.J. *et al.* (2002) Visual categorization and the primate prefrontal cortex: neurophysiology and behavior. *J. Neurophysiol.* 88, 929–941

196. Liu, T. (2016) Neural representation of object-specific attentional priority. *Neuroimage* 129, 15–24