

Top-down modulation: bridging selective attention and working memory

Adam Gazzaley¹ and Anna C. Nobre²

¹Departments of Neurology, Physiology and Psychiatry, University of California San Francisco, San Francisco, CA 94158, USA

²Department of Experimental Psychology and Oxford Centre for Human Brain Activity, University of Oxford, Oxford OX1 3UD, UK

Selective attention, the ability to focus our cognitive resources on information relevant to our goals, influences working memory (WM) performance. Indeed, attention and working memory are increasingly viewed as overlapping constructs. Here, we review recent evidence from human neurophysiological studies demonstrating that top-down modulation serves as a common neural mechanism underlying these two cognitive operations. The core features include activity modulation in stimulus-selective sensory cortices with concurrent engagement of prefrontal and parietal control regions that function as sources of top-down signals. Notably, top-down modulation is engaged during both stimulus-present and stimulus-absent stages of WM tasks; that is, expectation of an ensuing stimulus to be remembered, selection and encoding of stimuli, maintenance of relevant information in mind and memory retrieval.

Top-down modulation in perception and working memory

Selective attention and working memory (WM) have traditionally been viewed as distinct cognitive domains (see [Glossary](#)). However, a growing number of psychological and neuroscientific studies have revealed extensive overlap between these two constructs, spurring a number of excellent reviews and theoretical discussions on this topic [1–5]. In this review, we focus on recent neural evidence highlighting how top-down modulatory mechanisms, similar to those described for selective attention to stimuli for immediate perceptual goals, also influence multiple stages of representations that support WM performance.

Top-down modulation underlies our ability to focus attention on task-relevant stimuli and ignore irrelevant distractions. Investigations using an array of methodologies (see [Glossary](#)) have offered complementary contributions showing that top-down signals generate neural contrast by enhancing activity in sensory regions for items that are relevant and suppressing activity for items irrelevant to task goals. For example, in visual areas, excitability changes are reflected in competition for representations in receptive fields of individual neurons, baseline firing rates of neurons and synchronization of neuronal ensembles [6,7]. It is now widely accepted that top-down modulation

of sensory processing is not an intrinsic property of sensory cortices, but rather relies on long-range inputs from and interactions with a network of ‘control’ regions, including the prefrontal cortex (PFC) and parietal cortex [8,9]. We review evidence that a similar functional neural architecture of top-down modulation analogous to those that operate during perceptual analysis supports the prioritization of information in the service of WM.

In a typical visual WM task, participants are presented with an array of one or more items to be maintained in mind after the array is turned off over an interval of seconds (delay period) during which no stimulus information is present (‘delayed-response’ tasks). A single probe item or a probe array then appears, and the participant

Glossary

Selective attention: goal-directed focus on one aspect of the environment, while ignoring irrelevant aspects.

Working memory: maintenance and/or manipulation of task-relevant information in mind for brief periods of time to guide subsequent behavior.

Delay-response tasks: a commonly used cognitive paradigm to study working memory. It is particularly valuable for use in neural studies, because the stages of working memory can be dissociated in time: expectation, encoding, maintenance and retrieval. Types of delayed-response tasks include, delayed match to sample (DMTS), delayed non-match to sample (DNMTS), delayed-recognition tasks and the change-detection task.

Top-down modulation: modulation of neural activity in neurons in lower-order sensory or motor areas based on an individual’s goals. This may involve enhancement of task-relevant representations and/or suppression for task-irrelevant representations.

Functional connectivity: an analytical approach used to reveal how brain regions interact as nodes within neural networks and how their interactions change according to experimental variables; most frequently assessed as correlations between distributed neurophysiological measures in the time or time-frequency domain (coherence).

Functional MRI (fMRI): a whole-brain imaging technique that involves recording blood flow correlates of changes in neural activity during a task, often using the blood oxygen level-dependent (BOLD) signal. Spatial resolution is a strength (the order of millimeters), whereas temporal resolution is constrained by hemodynamic variables (the order of seconds) and is therefore a limitation.

Electroencephalography (EEG) and magnetoencephalography (MEG): non-invasive physiological recordings that sample ongoing neural electrical activity through sensors on the scalp. The advantage is high temporal resolution (on the order of milliseconds), but spatial resolution is limited due to spatial and temporal summation of neural activity within the brain volume. Data are most frequently analyzed as event-related potentials or magnetic (ERPs), which rely on averaging signals time-locked to stimulus processing in the time-domain, and spectral measures in the time-frequency domain.

Transcranial magnetic stimulation (TMS): non-invasive brain stimulation method that uses a rapidly changing magnetic field to induce an electrical current in underlying cortex. The unique strength is enabling functional assessment of brain regions and networks in a causal manner; a weakness is that accessibility to some cortical structures is limited.

Corresponding authors: Gazzaley, A. (adam.gazzaley@ucsf.edu); Nobre, A.C. (kia.nobre@ohba.ox.ac.uk).

judges whether the probe matches the item(s) of the previous array. Several stages of processing and neural representations occur in both the absence and presence of stimuli: a state of expectation can precede the display of to-be-remembered items; encoding of the items follows the presentation of the array; the delay period necessitates maintenance of the items in mind, in WM; and finally, presentation of the probe requires retrieval of the relevant item(s) from memory, as well as comparison to the probe, decision-making and responding. All of these stages are important in determining memory performance outcome, and all benefit from selective and focused processing. Furthermore, making only a subset of array items relevant to the task can intensify pressures for selective processing; that is, it introduces distraction. Selective processing can also be encouraged by providing instructive or predictive cues about the relevance of certain items for performance. Recent neural investigations manipulating selective pressures within each of these stages have revealed that top-down attentional modulatory mechanisms continue to operate throughout, dynamically regulating neuronal excitability to optimize the final WM performance outcome.

Expectation

Expectations of upcoming events generated by predictive cues enhance perceptual performance, notably improving the speed and accuracy of stimulus detection and discrimination [10]. The neural basis of this phenomenon has been most frequently studied using perceptual tasks, and is characterized by stimulus-absent, top-down modulation of neural activity in sensory cortices before stimulus presentation. For example, pre-stimulus enhancement has been demonstrated in visual cortices following predictive cues that selectively guide attention to a location [11], stimulus feature [12,13], or object [14–16]. As is true for top-down modulation in the presence of a stimulus, the prevailing view is that top-down modulation of pre-stimulus activity is mediated via top-down signals from the PFC and parietal control regions [11,17,18].

Predictive cues have now also been shown to aid in the transfer of perceptual representations into WM, as illustrated by the ability of such cues to benefit WM performance [19–23]. A recent functional magnetic resonance imaging (fMRI) study exploring the neural basis of this phenomenon documented top-down modulation of pre-encoding activity as a mechanism for expectation-driven WM benefits [24]. Bollinger *et al.* used an object delayed-response task in which a cue predicted what stimulus category would be presented at encoding (e.g. predictive = face, neutral = face or scene). Predictive category cueing led to enhanced WM for faces compared to faces following a neutral cue, and triggered shifts in baseline activity in a face-processing region of the visual association cortex (i.e. the fusiform face area [FFA]). The degree of functional connectivity between the FFA and regions in the PFC and parietal cortex (right inferior frontal junction [IFJ], middle frontal gyrus [MFG], inferior frontal gyrus [IFG] and intraparietal sulcus [IPS]) correlated with the magnitude of pre-stimulus activity modulation in the FFA. Moreover, FFA functional connectivity with the IFJ predicted the benefit in WM performance gained by expectation of a

specific stimulus category. These data support the role of stimulus-absent, top-down modulation in mediating the influence of expectations on WM performance. Complementing these findings, Murray *et al.* demonstrated biasing of pre-encoding activity by spatial attention and a correlation of these biases with WM performance using electroencephalography (EEG) [21] (Box 1).

McNab *et al.* [25] used a feature delayed-response task that cued participants to the presence or absence of irrelevant information in the encoding stimulus set. fMRI data attributed to the pre-encoding expectation period revealed greater activity in the PFC (MFG, near the precentral sulcus) and the left basal ganglia (putamen and global pallidus) when participants were cued to expect distraction. The authors interpret these brain regions as participating in establishing a ‘filtering set’ to focus encoding resources only on the relevant information. Consistent with this, we suggest this increased activity reflects the greater requirement for selective attention to the relevant stimulus set in the context of distracting information.

Encoding

The differential modulation of activity in sensory cortices by attention to relevant versus irrelevant stimuli during WM encoding has been shown to be comparable to activity modulation generated in purely perceptual tasks. Goal-related influences occur at both early and late phases of stimulus processing [26] in stimulus-selective sensory cortices [27,28]. There is now accumulating support for a direct relationship between early goal-driven activity modulation in sensory areas (within 200 ms of stimulus onset) and subsequent WM performance [29]. For example, Rutman

Box 1. Cognitive aging

It is well established that older adults experience deficits in both attention and WM abilities [89]. It is becoming increasingly evident that these deficits interact with one another. In the context of this review, there is accumulating evidence that impaired selective attention processes in aging underlie much of the WM deficits experienced by older adults. This is documented by alterations in neural markers of top-down modulation throughout the stages of a delayed-response task. In terms of expectation, unlike younger individuals, older adults exhibit deficits in using predictive information to guide optimal perceptual performance (e.g. temporal attention [90]). This age-related expectation deficit has now been associated with an inability of older adults to achieve WM benefits using predictive cues, as do younger adults [24,91]. fMRI evidence further reveals that alterations in top-down control networks underlie an absence of pre-encoding, visual cortical activity modulation and associated WM benefits in older adults. Older adults also do not engage top-down modulation mechanisms during WM encoding in the setting of irrelevant information to the same degree as younger adults. This age-related impairment in top-down modulation is selective for deficits in the suppression of distracting information [92] and occurs at early visual processing stages [93], even if an older individual can anticipate the presentation of a distractor [94,95]. In a related manner, older adults experience a greater impact on memory when multitasking during a period of WM maintenance compared to younger adults. This is associated with a neural deficit in effectively disengaging from interruptions and re-establishing functional connectivity associated with the memory network after interruption [96]. As a whole, these findings converge to reveal a generalizable, age-related deficit in top-down modulation that serves as a bridge between attention and WM impairment in aging [97].

et al. [30] used EEG and a delayed-response task with overlapping face and scene stimuli to reveal that activity modulation driven by selective attention occurs within 100 ms of stimulus encoding and predicts subsequent WM accuracy (Figure 1). The interpretation was that early modulation of cortical activity minimizes interference from distractors and biases the generation of higher fidelity representations of relevant stimuli, thus conferring an advantage in maintaining that information in mind. A similar study that used a feature delayed-response task [31] further revealed that optimal WM performance was dependent on effectively filtering irrelevant information at early processing stages, presumably to prevent overloading of a limited WM capacity, supporting research by Vogel and colleagues [32].

fMRI studies assessing cortical control regions involved in top-down modulation during WM encoding have converged to reveal a role of the PFC both in processing relevant stimuli and in filtering distractors. Functional connectivity studies showed that a region in the PFC (left MFG) was more strongly functionally connected with a scene-processing visual region when scenes were remembered and less so when scenes were ignored during the encoding phase of an object delayed-response task [33]. Moreover, the strength of functional connectivity correlated with the magnitude of attentional enhancement for relevant stimuli and suppression of irrelevant stimuli, suggesting that the PFC modulates activity levels in visual cortices via the strength of functional coupling. Interestingly, a recent fMRI study using an object delayed-response task with overlapping stimuli revealed that visual cortical areas processing relevant object information were functionally connected with the PFC and parietal control areas (i.e. MFG, bilateral IFJ and IPS), whereas visual cortical areas processing irrelevant stimuli were simultaneously coupled with the 'default network' (i.e. medial PFC and posterior cingulate cortex) [34]. Importantly, there was a relationship between WM performance and the degree of coupling between visual cortices and default network regions.

These studies support a role of the PFC in selective attention processes engaged during WM encoding, and raise the possibility that limitations in attentional allocation by

the PFC may serve as a limiting factor in the amount of information we can encode. An fMRI study explicitly assessed common neural resources shared by selective attention for perceptual goals and WM encoding by combining a visual search and object delayed-response task along with manipulations of task demands [35]. The results revealed overlapping activations for search and WM encoding in prefrontal, visual, parietal and premotor areas. Moreover, there was a reduction in WM load response in the setting of high attentional demand, suggesting that shared neural resources between attention and WM encoding limit processing capabilities.

Despite the informative contributions made by fMRI studies, they generate correlational data. To test the causal role of PFC-mediated top-down modulation on subsequent WM performance, a recent study used fMRI-guided, repetitive transcranial magnetic stimulation (rTMS) to perturb function within a PFC region, and followed this with EEG recordings during a feature delayed-response task [36]. The right IFJ region targeted for rTMS was previously identified as a putative PFC control node in an fMRI study [37]. Coordinates for the IFJ rTMS were based on functional connectivity analysis of fMRI data for each participant using the same WM task. Ten minutes of rTMS to the right IFJ resulted in significantly diminished top-down modulation of the P1 component of the event-related potential (ERP) to color stimuli at posterior electrodes (i.e. the difference between P1 amplitude time-locked to the onset of relevant and irrelevant fields of colored dots), as well as a significant reduction in WM accuracy for color. As P1 modulation recovered with time after rTMS (i.e. in the second half of the block), so did WM performance. Moreover, on an individual participant basis, the rTMS-induced reduction in P1 modulation during color processing predicted the reduction in color WM accuracy. This study further showed that participants with stronger fMRI functional connectivity between the IFJ and visual cortices displayed a greater impact of IFJ rTMS on top-down modulation. Thus, the results revealed that PFC-mediated top-down modulation during the early visual processing stages of WM encoding was causally related to subsequent WM performance (Figure 2). The right IFJ was identified as a PFC control

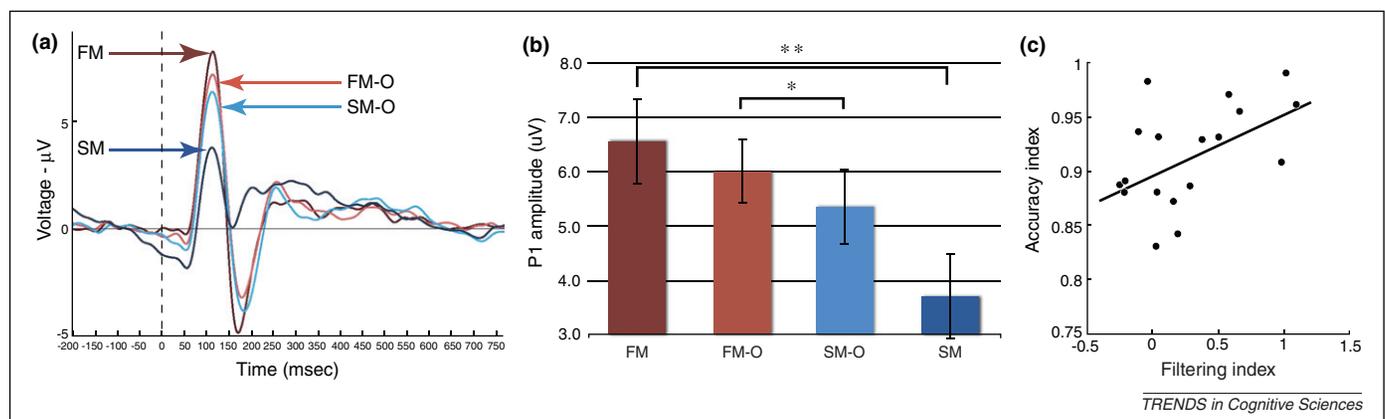


Figure 1. Correlation between WM performance and early neural measures of selective activity modulation in the visual cortices during encoding. (a) Grand-average ERP waveform showing early activity modulation and (b) significant differences in P1 amplitude based solely on WM goals (FM-O versus SM-O). (c) Participants with greater attentional modulation of P1 amplitude (approximately 100 ms post-stimulus presentation) show greater subsequent memory of encoded stimuli ($R = 0.45$, $p < 0.05$). FM-O, face memory-overlap; SM-O, scene memory-overlap; FM, face memory; SM, scene memory. Error bars represent the standard error of the mean. Asterisks denote significant differences (single $p < 0.05$, double $p < 0.001$). Modified from [30].

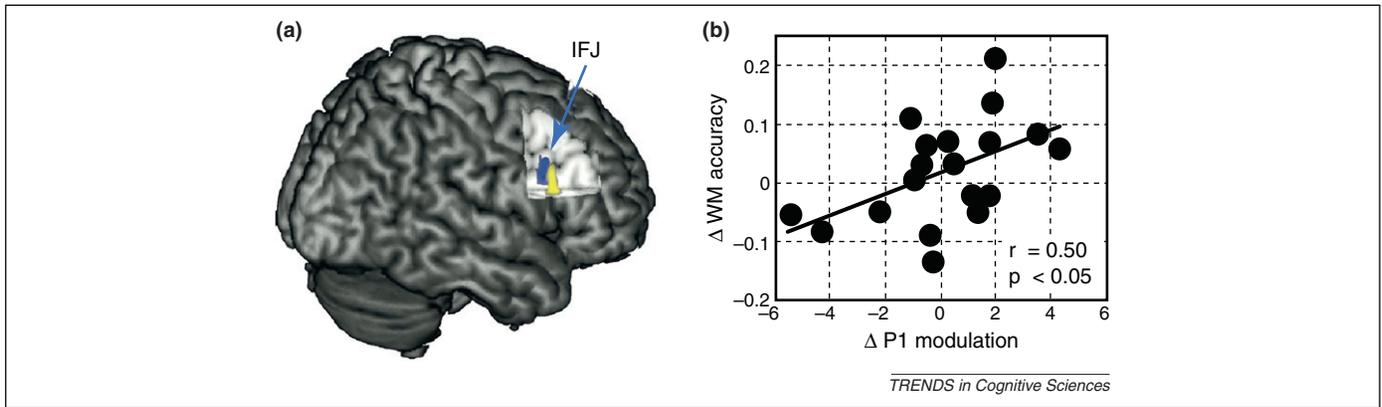


Figure 2. Site of rTMS and relationship between rTMS-induced changes in P1 modulation and WM accuracy. (a) Functional connectivity analysis revealed a PFC region associated with top-down modulation, the right inferior frontal junction (IFJ), that was identified in each individual and served as the site for rTMS. (b) Participants with greater rTMS-related change in P1 modulation also exhibited a greater change in WM accuracy. Δ = sham – actual rTMS. Modified from [36].

region that mediated this causal connection between top-down modulation in the service of attentional goals and WM performance.

Maintenance

Encoding into WM was traditionally considered to be the end point of selective attention. Early seminal work showed that retrieval of items from a previously presented array could be improved only by selective spatial retrieval cues (post-cues) appearing after very brief intervals, when representations were thought to exist in a rapidly decaying photographic ‘iconic’ format. Post-cues became ineffective shortly afterward, during WM maintenance [38–40]. Once encoded, WM representations were thought to be stable, resistant to interference, and accessible through an automatic and serial-search mechanism [41].

Subsequent research instead promoted the view that selective attention continues to operate during WM maintenance. One prevailing view is that selective attention is the mechanism by which items remain or become activated, and thus are maintained, in WM [2–5]. Corroborating evidence comes from behavioral and neural studies showing improved visual processing of items appearing at the same location as or sharing features with items being actively maintained in WM [42–47]. According to this view, the functional architecture that supports visual WM maintenance is analogous to that supporting visual selective attention in the perceptual domain [3,48,49]. fMRI task-based functional connectivity suggests that maintenance of items in WM involves sustained interactions between the perceptual areas coding the relevant attributes of the items and control areas in the PFC and parietal cortices [50]. Parietal and frontal areas may therefore play an attention-directing role during WM maintenance [51,52], a view supported by a recent study investigating deficits in WM tasks after bilateral lesions to the posterior parietal cortex (PPC) [61].

Current research advances an even more active role for top-down modulation during WM maintenance. In line with the changing task demands and expectations in a dynamic environment, top-down signals can continue to prioritize items being maintained in WM even after encoding [1,53,54]. An important methodological tool has been the introduction of ‘retro-cues’ during the WM maintenance

period to manipulate retroactively knowledge or expectations about which maintained items will be relevant for subsequent behavior. Retro-cues differ from post-cues in that they do not prompt immediate retrieval of a cued item, but instead provide information about the relevance of a given item for subsequent retrieval. They trigger top-down biasing mechanisms that operate on representations being maintained in WM and consistently confer large benefits to WM performance [23,55,56]. Recent neural findings suggest that the mechanisms of stimulus-absent, top-down modulation during WM maintenance are similar to mechanisms for attentional modulation during perception, but may also involve additional regulatory functions.

Retro-cues directing attention to a given category of object within visual WM have been shown dynamically to modulate activity in visual areas coding relevant versus irrelevant items [57,58]. Similar activation is also observed in tasks requiring participants merely to think back to a previously viewed item; that is, to ‘refresh’ their current focus of WM [59,60]. Spatial retro-cues can also result in retinotopically specific modulation of visual activity during WM maintenance [61]. The pattern of modulation is remarkably similar to that of anticipatory selective attention during expectation periods in perceptual [11,15,16,14] and WM tasks [24]. Although these results invite the conclusion that common top-down mechanisms bias neuronal excitability toward an expected percept or toward a maintained memorandum, an alternative possibility is that the modulation of activity after retro-cues or refreshing cues is unrelated to modulation of WM, and merely reflects anticipatory attention toward the expected probe stimulus [62]. Kuo and colleagues [63] recently used spatial retro-cues in a delayed-response task that eliminated any anticipatory spatial attention to the probe. The data, which relied on lateralized ERPs reflecting load-dependent maintenance of visual activity (contralateral delay activity [CDA] [64]), provided strong support for the ability to dynamically modulate WM representations themselves during the maintenance period.

fMRI studies isolating neural activity triggered by informative retro-cues show that the PFC and parietal areas participate in controlling the focus of attention among items being maintained in WM, when orienting attention to both locations [63,65,66] and objects [57,62,67,68] in

delayed-response tasks. Similar activation is observed after 'refresh' cues [69,70]. Direct comparisons between cues orienting attention within WM versus within perceptual arrays show a common set of parietal and PFC control areas, with some studies showing additional recruitment of medial and lateral PFC areas when the focus of attention within WM is controlled [65,67,71,72]. fMRI functional connectivity between the PFC and visual areas increases after effective spatial retro-cues, and correlates significantly with WM performance measures in a feature delayed-response task [73]. Analyses using multi-voxel pattern classification in a recent study showed that shifting the focus of attention between two perceptual streams versus two mental counters activated different patterns of neurons within common parietal and frontal areas [72]. Further studies using such multivariate methods will be informative in revealing whether there are systematic mappings within parietal and frontal 'control' areas according to the domain in which attention shifts, or whether neurons become bootstrapped into participating in top-down modulation according to a flexible, adaptive coding mechanism [74,75].

The left ventrolateral PFC, around the IFG and IFJ, has been particularly implicated in regulating excitability among WM representations [57,63,65–69,76]. A recent study combining rTMS and fMRI supports the causal participation of this general area in focusing attention during WM maintenance and regulating neural excitability in visual areas accordingly [58]. A region in the left ventrolateral cortex, therefore, may play a key control function related to selecting from among memory representations to guide perception and action [58,67,68,77].

Another frequently used task manipulation to study selective attention processes during WM maintenance is the introduction of distracting stimuli or a secondary task (interruption) during the delay period. Distractions presented during maintenance diminish WM task performance, especially when distractors are of the same category as memoranda [78–81]. Accumulating evidence suggests that selective attention biases sensory processing in favor of the information being retained in WM and against irrelevant, distracting information in a manner comparable to that during perceptual processing [79,81]. Moreover, ERPs and fMRI have shown that the degree of visual processing of irrelevant, distracting stimuli inversely correlates with subsequent WM recognition accuracy for the encoded items [78,82]. fMRI connectivity analyses suggests that, in the presence of entirely irrelevant distracting stimuli, encoded items are maintained throughout the delay period via retained functional connectivity between the MFG and visual areas. In contrast, when maintenance was interrupted by a secondary task, the encoded items were not maintained but rather reactivated after the interruption [78]. A recent combined TMS–fMRI study showed that the strength of top-down inputs from the dorsolateral prefrontal cortex to posterior brain areas processing task-relevant targets during WM maintenance increased in the presence of distracting stimuli, confirming the importance of functional connectivity between the PFC and posterior areas in promoting successful maintenance of task-relevant items during WM delays [83].

Retrieval

Given the multiple stages of the delayed-recognition task at which top-down modulation can influence WM performance, finally we must ask whether the retrieval process itself is influenced by these mechanisms. This issue seems to have undergone little direct investigation. Conceptual as well as pragmatic issues may make the question difficult to tackle. For example, it is difficult to separate effects operating at retrieval from the cascading consequences of effects that have accrued during expectation, encoding and maintenance. Nobre and colleagues [84] measured ERPs while participants searched for a target in a WM array of varying loads (1–4). In the absence of spatial cueing, the probe triggered a load-dependent ERP reflecting WM search [85,86]; spatial retro-cues completely abolished this load-related activity. These results clearly show that selective attention can facilitate retrieval functions, but leaves it unclear whether the effect occurred at the time of retrieval (e.g. by using a short-cut to the relevant location in the search path) or earlier during the maintenance period (e.g. by effectively reducing the relevant array to one item), or both. Some ingenuity will be required to separate out independent modulatory effects during WM retrieval.

Recent studies of WM retrieval have revealed another type of close connection between visual WM and attention. Selective retrieval of a target item maintained within a WM array elicits lateralized ERP markers similar to those obtained during successful identification of a target during visual search [85–88]. The lateralized potentials occur even when the WM search is triggered by a centrally presented target, and are thought to reflect activation of WM representations of target attributes within a spatiotopically organized layout in visual and/or parietal areas [86]. Thus, selecting a target from WM and from the environment may be largely analogous acts of internally and externally directed selective attention.

Concluding remarks

We conclude that the neural mechanism of top-down modulation serves as a common framework for selective attention processes in the service of both perceptual goals and those that underlie the different stages of WM operations (Box 2). Top-down modulation of information processing appears analogous between these constructs, in terms of both the site of activity modulation in the sensory cortices and the putative sources of top-down signals originating from cortical control areas. We propose that the role of top-down modulation during WM encoding and its subsequent benefits on WM performance is via an influence of attention on the early perceptual representation, and is not a process specific for WM. This also applies to the modulation of sensory representations that occur before stimulus presentation and during the delay period. Top-down modulatory functions dynamically modulate neuronal excitability both in the presence of stimuli (i.e. during selective encoding of items to be remembered and selective retrieval of a memorandum) and in the absence of external stimuli (i.e. in expectation of items to encode or ignore and during maintenance of items during a temporal delay).

Box 2. Questions for future research

- Given the striking similarities between the neural mechanisms engaged when selecting and representing items from an incoming perceptual stream and those that are engaged in the absence of stimulation, one wonders exactly how percepts and WM traces are distinguished at a neural level.
- The modulation of neural oscillations has increasingly been shown to play a critical role in regulating excitability for perceptual events [7], but we know little about the role of oscillations in the top-down modulation of information maintained in WM. Studies should explore phase and amplitude relationships at different frequencies, both within and across brain regions, and across the stages of WM tasks.
- Goal-directed focus on relevant events is a core building block for most of cognition, and disturbances in attention have been implicated in several neuropsychiatric conditions. It will be of great interest to explore whether and how alterations in top-down modulation at the interface of attention and WM impact cognitive performance in neuropsychiatric conditions.
- How amenable are these neural systems to plasticity changes with training? Is it possible to enhance cognitive development during childhood and resilience during aging by improving focus and selective attention within WM?

Acknowledgements

This work was supported by the National Institutes of Health R01-AG030395 (AG), the American Federation of Aging Research (AG), the Ellison Medical Foundation (AG), the James S. McDonnell Foundation (ACN) and the Wellcome Trust (ACN).

References

- Chun, M.M. (2011) Visual working memory as visual attention sustained internally over time. *Neuropsychologia* 49, 1407–1409
- Cowan, N. (1995) *Attention and Memory: An Integrated Framework*, Oxford University Press
- Postle, B.R. (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38
- Awh, E. and Jonides, J. (2001) Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.* 5, 119–126
- Awh, E. et al. (2006) Interactions between attention and working memory. *Neuroscience* 139, 201–208
- Reynolds, J.H. and Chelazzi, L. (2004) Attentional modulation of visual processing. *Annu. Rev. Neurosci.* 27, 611–647
- Fries, P. (2009) Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* 32, 209–224
- Curtis, C.E. and D'Esposito, M. (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423
- Gazzaley, A. and D'Esposito, M. (2007) Unifying prefrontal cortex function: executive control, neural networks and top-down modulation, In *The Human Frontal Lobes* (2nd edn) (Cummings, J. and Miller, B., eds), pp. 187–206, Guilford Press
- Posner, M.I. (1980) Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25
- Kastner, S. et al. (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761
- Chawla, D. et al. (1999) The physiological basis of attentional modulation in extrastriate visual areas. *Nat. Neurosci.* 2, 671–676
- Giesbrecht, B. et al. (2006) Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Res.* 1080, 63–72
- Esterman, M. and Yantis, S. (2010) Perceptual expectation evokes category-selective cortical activity. *Cereb. Cortex* 20, 1245–1253
- Puri, A. et al. (2009) Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Res.* 1301, 89–99
- Stokes, M. et al. (2009) Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19569–19574
- Corbetta, M. and Shulman, G.L. (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215
- Bressler, S.L. et al. (2008) Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *J. Neurosci.* 28, 10056–10061
- Palmer, J. (1990) Attentional limits on the perception and memory of visual information. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 332–350
- Schmidt, B. et al. (2002) Voluntary and automatic attentional control of visual working memory. *Percept. Psychophys.* 64, 754–763
- Murray, A.M. et al. (2011) Markers of preparatory attention predict visual short-term memory performance. *Neuropsychologia* 49, 1458–1465
- Botta, F. et al. (2010) Exogenous and endogenous spatial attention effects on visuospatial working memory. *Q. J. Exp. Psychol. (Colchester)* 63, 1590–1602
- Griffin, I.C. and Nobre, A.C. (2003) Orienting attention to locations in internal representations. *J. Cogn. Neurosci.* 15, 1176–1194
- Bollinger, J. et al. (2010) Expectation-driven changes in cortical functional connectivity influence working memory and long-term memory performance. *J. Neurosci.* 30, 14399–14410
- McNab, F. and Klingberg, T. (2008) Prefrontal cortex and basal ganglia control access to working memory. *Nat. Neurosci.* 11, 103–107
- Hillyard, S.A. et al. (1998) Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 353, 1257–1270
- Gazzaley, A. et al. (2005) Top-down enhancement and suppression of the magnitude and speed of neural activity. *J. Cogn. Neurosci.* 17, 507–517
- Kastner, S. et al. (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282, 108–111
- Gazzaley, A. (2011) Influence of early attentional modulation on working memory. *Neuropsychologia* 49, 1410–1424
- Rutman, A.M. et al. (2010) Early top-down control of visual processing predicts working memory performance. *J. Cogn. Neurosci.* 22, 1224–1234
- Zanto, T.P. and Gazzaley, A. (2009) Neural suppression of irrelevant information underlies optimal working memory performance. *J. Neurosci.* 29, 3059–3066
- Vogel, E.K. et al. (2005) Neural measures reveal individual differences in controlling access to working memory. *Nature* 438, 500–503
- Gazzaley, A. et al. (2007) Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. *Cereb. Cortex* 17 (Suppl. 1), i125–i135
- Chadick, J.Z. and Gazzaley, A. (2011) Differential coupling of visual cortex with default or frontal-parietal network based on goals. *Nat. Neurosci.* 14, 830–832
- Mayer, J.S. et al. (2007) Common neural substrates for visual working memory and attention. *Neuroimage* 36, 441–453
- Zanto, T.P. et al. (2011) Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nat. Neurosci.* 14, 656–661
- Zanto, T.P. et al. (2010) Top-down modulation of visual feature processing: the role of the inferior frontal junction. *Neuroimage* 53, 736–745
- Sperling, G. (1960) The information available in brief visual presentations. *Psychol. Monogr.* 74, 1–29
- Averbach, E. and Coriell, A.S. (1961) Short-term memory in vision. *Bell Syst. Tech. J.* 40, 309–328
- Phillips, W.A. (1974) On the distinction between sensory storage and short term visual memory. *Percept. Psychophys.* 16, 283–290
- Sternberg, S. (1966) High-speed scanning in human memory. *Science* 153, 652–654
- Awh, E. et al. (2000) The role of spatial selective attention in working memory for locations: evidence from event-related potentials. *J. Cogn. Neurosci.* 12, 840–847
- Awh, E. et al. (1998) Rehearsal in spatial working memory. *Percept. Perform.* 24, 780–790
- Downing, P.E. (2000) Interactions between visual working memory and selective attention. *Psychol. Sci.* 11, 467–473
- Jha, A.P. (2002) Tracking the time-course of attentional involvement in spatial working memory: an event-related potential investigation. *Brain Res. Cogn. Brain Res.* 15, 61–69
- Soto, D. et al. (2008) Automatic guidance of attention from working memory. *Trends Cogn. Sci.* 12, 342–348
- Postle, B. et al. (2004) The where and how of attention-based rehearsal in spatial working memory. *Cogn. Brain Res.* 20, 194–205

- 48 McCarthy, G. (1995) Functional neuroimaging of memory. *Neuroscientist* 1, 155–163
- 49 Lepsien, J. and Nobre, A.C. (2006) Cognitive control of attention in the human brain: insights from orienting attention to mental representations. *Brain Res.* 1105, 20–31
- 50 Gazzaley, A. *et al.* (2004) Functional connectivity during working memory maintenance. *Cogn. Affect. Behav. Neurosci.* 4, 580–599
- 51 Magen, H. *et al.* (2009) Attentional demands predict short-term memory load response in posterior parietal cortex. *Neuropsychologia* 47, 1790–1798
- 52 Mitchell, D. *et al.* (2008) A review of frontal-midline theta from the perspective of hippocampal “THETA”. *Prog. Neurobiol.* 86, 156–185
- 53 Murty, V.P. *et al.* (2011) Selective updating of working memory content modulates meso-cortico-striatal activity. *Neuroimage* 57, 1264–1272
- 54 Stokes, M. and Nobre, A.C. (2011) Top-down biases in visual short-term memory. In *Neuroscience of Attention: Attentional Control and Selection* (Mangun, G.R., ed.), pp. 209–228, Oxford University Press
- 55 Sligte, I.G. *et al.* (2008) Are there multiple visual short-term memory stores? *PLoS ONE* 3, e1699
- 56 Landman, R. *et al.* (2003) Large capacity storage of integrated objects before change blindness. *Vision Res.* 43, 149–164
- 57 Lepsien, J. and Nobre, A.C. (2007) Attentional modulation of object representations in working memory. *Cereb. Cortex* 17, 2072–2083
- 58 Higo, T. *et al.* (2011) Distributed and causal influence of frontal operculum in task control. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4230–4235
- 59 Johnson, M.R. and Johnson, M.K. (2009) Top-down enhancement and suppression of activity in category-selective extrastriate cortex from an act of reflective attention. *J. Cogn. Neurosci.* 21, 2320–2327
- 60 Johnson, M.R. *et al.* (2007) A brief thought can modulate activity in extrastriate visual areas: top-down effects of refreshing just-seen visual stimuli. *Neuroimage* 37, 290–299
- 61 Sligte, I.G. *et al.* (2009) V4 activity predicts the strength of visual short-term memory representations. *J. Neurosci.* 29, 7432–7438
- 62 Lepsien, J. *et al.* (2011) Modulation of working-memory maintenance by directed attention. *Neuropsychologia* 49, 1569–1577
- 63 Kuo, B.C. *et al.* (2012) Attention modulates maintenance of representations in visual short-term memory. *J. Cogn. Neurosci.* 24, 51–60
- 64 Vogel, E.K. and Machizawa, M.G. (2004) Neural activity predicts individual differences in visual working memory capacity. *Nature* 428, 748–751
- 65 Nobre, A.C. *et al.* (2004) Orienting attention to locations in perceptual versus mental representations. *J. Cogn. Neurosci.* 16, 363–373
- 66 Lepsien, J. *et al.* (2005) Directing spatial attention in mental representations: interactions between attentional orienting and working-memory load. *Neuroimage* 26, 733–743
- 67 Nee, D.E. and Jonides, J. (2009) Common and distinct neural correlates of perceptual and memorial selection. *Neuroimage* 45, 963–975
- 68 Roth, J.K. *et al.* (2006) Neural system for controlling the contents of object working memory in humans. *Cereb. Cortex* 16, 1595–1603
- 69 Raye, C.L. *et al.* (2007) Refreshing: a minimal executive function. *Cortex* 43, 135–145
- 70 Roth, J.K. and Courtney, S.M. (2007) Neural system for updating object working memory from different sources: sensory stimuli or long-term memory. *Neuroimage* 38, 617–630
- 71 Bledowski, C. *et al.* (2009) What “works” in working memory? Separate systems for selection and updating of critical information. *J. Neurosci.* 29, 13735–13741
- 72 Tamber-Rosenau, B.J. *et al.* (2011) Cortical mechanisms of cognitive control for shifting attention in vision and working memory. *J. Cogn. Neurosci.* 23, 2905–2919
- 73 Kuo, B.C. *et al.* (2011) Functional connectivity during top-down modulation of visual short-term memory representations. *Neuropsychologia* 49, 1589–1596
- 74 Duncan, J. (2001) An adaptive coding model of neural function in the prefrontal cortex. *Nat. Rev. Neurosci.* 2, 820–829
- 75 Woolgar, A. *et al.* (2011) Adaptive coding of task-relevant information in human frontoparietal cortex. *J. Neurosci.* 31, 14592–14599
- 76 Johnson, M.K. *et al.* (2005) Using fMRI to investigate a component process of reflection: prefrontal correlates of refreshing a just-activated representation. *Cogn. Affect. Behav. Neurosci.* 5, 339–361
- 77 Curtis, C. and D’Esposito, M. (2003) Success and failure suppressing reflexive behavior. *J. Cogn. Neurosci.* 15, 409–418
- 78 Clapp, W.C. *et al.* (2010) Mechanisms of working memory disruption by external interference. *Cereb. Cortex* 20, 859–872
- 79 Dolcos, F. *et al.* (2007) Regional brain differences in the effect of distraction during the delay interval of a working memory task. *Brain Res.* 1152, 171–181
- 80 Yoon, J.H. *et al.* (2006) Differential effects of distraction during working memory on delay-period activity in the prefrontal cortex and the visual association cortex. *Neuroimage* 29, 1117–1126
- 81 Sreenivasan, K.K. and Jha, A.P. (2007) Selective attention supports working memory maintenance by modulating perceptual processing of distractors. *J. Cogn. Neurosci.* 19, 32–41
- 82 Berry, A.S. *et al.* (2009) Practice-related improvement in working memory is modulated by changes in processing external interference. *J. Neurophysiol.* 102, 1779–1789
- 83 Feredoes, E. *et al.* (2011) Causal evidence for frontal involvement in memory target maintenance by posterior brain areas during distracter interference of visual working memory. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17510–17515
- 84 Nobre, A.C. *et al.* (2007) Spatial attention can bias search in visual short-term memory. *Front. Hum. Neurosci.* 1, 4
- 85 Leszczyński, M. *et al.* (2011) Recoding between two types of STM representation revealed by the dynamics of memory search. *J. Cogn. Neurosci.* DOI: 10.1162/jocn_a_00102
- 86 Kuo, B.C. *et al.* (2009) Searching for targets within the spatial layout of visual short-term memory. *J. Neurosci.* 29, 8032–8038
- 87 Dell’Acqua, R. *et al.* (2010) Orienting attention to objects in visual short-term memory. *Neuropsychologia* 48, 419–428
- 88 Eimer, M. and Kiss, M. (2010) An electrophysiological measure of access to representations in visual working memory. *Psychophysiology* 47, 197–200
- 89 Craik, F.I. and Salthouse, T.A. (2000) *Handbook of Aging and Cognition II*, Erlbaum
- 90 Zanto, T.P. *et al.* (2011) Age-related changes in orienting attention in time. *J. Neurosci.* 31, 12461–12470
- 91 Bollinger, J. *et al.* (2011) An expectation-based memory deficit in aging. *Neuropsychologia* 49, 1466–1475
- 92 Gazzaley, A. *et al.* (2005) Top-down suppression deficit underlies working memory impairment in normal aging. *Nat. Neurosci.* 8, 1298–1300
- 93 Gazzaley, A. *et al.* (2008) Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13122–13126
- 94 Clapp, W.C. and Gazzaley, A. (2012) Distinct mechanisms for the impact of distraction and interruption on working memory in aging. *Neurobiol. Aging* 33, 134–148
- 95 Zanto, T.P. *et al.* (2010) Predictive knowledge of stimulus relevance does not influence top-down suppression of irrelevant information in older adults. *Cortex* 46, 564–574
- 96 Clapp, W.C. *et al.* (2011) A deficit in switching between functional brain networks underlies the impact of multitasking on working memory in older adults. *Proc. Natl. Acad. Sci. U.S.A.* DOI: 10.1073/pnas.1015297108
- 97 Gazzaley, A. (2011) Top-down modulation and cognitive aging. In *Principles of Frontal Lobe Function* (2nd edn) (Knight, R.T. and Stuss, D.T., eds), Oxford University Press (in press)