Short Term and Working Memory

B R Postle, University of Wisconsin – Madison, Madison, WI, USA
T Pasternak, University of Rochester, Rochester, NY, USA

Introduction

Definition

Short-term memory (STM) refers to the active retention (for humans, the 'keeping in mind') of information when it is not accessible from the environment. Working memory can be thought of as 'STM+,' the '+' referring to the ability to manipulate or otherwise transform this information, to protect it in the face of interference, and to use it in the service of such high-level behaviors as planning, reasoning, and problem solving. STM and working memory are of central importance to the study of high-level cognition because they are believed to be critical contributors to such essential cognitive functions and properties as language comprehension, learning, planning, reasoning, and general fluid intelligence.

Historical Backdrop

The modern study of working memory began with the work of Jacobsen, who, in the 1930s, demonstrated that large bilateral lesions of the prefrontal cortex (PFC) in the monkey produced profound deficits in spatial working memory, and thus Jacobsen ascribed to the PFC a function he termed "immediate memory." A subsequent study, however, revealed that this deficit could be erased by darkening the testing cage and thus reducing interference during the memory delay. These results highlighted a critical question, still the subject of an active debate, which is whether PFC plays a critical role in storage in addition to its role in behavioral control.

The early 1970s witnessed two developments that were seminal in shaping contemporary conceptions of working memory. The first was the observation made by Joaquin Fuster that individual neurons in PFC of monkeys, trained to compare visual stimuli separated by a memory delay, exhibit sustained activity throughout that delay. This observation suggested a neural correlate of two potent ideas from physiological psychology – that of a PFC-dependent immediate memory and that of a reverberatory mechanism for "a transient 'memory'" proposed by psychologist Donald Hebb. The second development, which occurred in the field of human cognitive psychology, was the multiple component model of working memory proposed by Alan Baddeley and Graham Hitch. In its initial instantiation this model comprised two independent buffers for the storage of verbal and of visuospatial information, and a central executive to control attention and to manage information in the buffers. Prompted by these two developments, the neuroscientific and the psychological study of working memory each proceeded along parallel, but largely independent, paths until the late 1980s, when a third important advance occurred.

The third advance was the conceptual integration of the neuroscientific and psychological traditions of working memory research, proposed by Patricia Goldman-Rakic, that the sustained delay-period neural activity in PFC and the storage buffers of the multiple-component model of Baddeley and Hitch were cross-species manifestations of the same fundamental mental phenomenon. This association between prefrontal cortex and working memory has been very influential in systems and cognitive neuroscience.

Current State of Working Memory Research

A growing body of evidence provided by behavioral, physiological, and neuroimaging studies indicates that information about sensory stimuli may be stored in a segregated, feature-selective manner, and that the relevant cortical regions include relatively early stages of sensory cortical processing. The principle emerging from this work is that the same brain regions that are responsible for the precise sensory encoding of information also contribute to its short-term retention. In the remaining portions of this article we briefly describe these recent advances, with the emphasis on working memory for fundamental dimensions of sensory stimuli.

Visual Working Memory

In the laboratory, the ability to briefly retain visual information can be measured with delayed discrimination tasks. In the simplest version of such tasks, individuals discriminate between two stimuli, the sample and the test, separated by a temporal delay of various durations, and report whether and how the test differs from the previously seen sample. The results of such experiments have revealed that fundamental stimulus features, such as orientation, contrast, size, or speed, can be faithfully preserved for many seconds, although the duration of this preservation often differs for different features. For example, stimulus size or orientation can be retained accurately longer than can luminance contrast. This

difference suggests that different stimulus attributes may be retained by separate, feature-selective mechanisms. More direct support for feature-selective storage mechanisms comes from studies that have used interference consisting of an irrelevant stimulus (a 'memory mask') introduced during the delay separating S1 and S2. With this procedure, determination of the parameters that maximize the interfering effects of the mask can provide insights into the nature of the remembered stimulus and thus into the mechanisms involved in its short-term retention. For example, the memory for spatial frequency of gratings (size) can be disrupted only if the masking stimulus is of a different spatial frequency, irrespective of its orientation. This selective interference suggests that information about spatial frequencies may be preserved by mechanisms that are narrowly tuned for that stimulus attribute, and closely associated with its processing, but distinct from mechanisms concerned with stimulus orientation. Similar specialization is seen within the domain of visual motion, in that stimulus speed can be preserved by the mechanisms that are relatively narrowly tuned for speed, independently of stimulus direction.

Representation of visual motion in memory appears to be localized in space in a manner consistent with properties of neurons in cortical area MT, a region specialized in processing of visual motion, thereby supporting the idea that neurons processing visual motion may also be involved in its short-term storage. Overall, behavioral studies of mechanisms that preserve basic attributes of visual stimuli can be characterized as narrowly tuned, spatially localized filters, supporting a model of working memory that involves the contribution of sensory cortical areas.

A close link between visual processing and storage is also suggested by the consequences of cortical damage. For example, damage to motion processing area MT in monkeys results in deficits in remembering motion direction, but only when the remembered stimuli require area MT for their encoding (Figure 1). Similarly, damage to inferotemporal cortex, associated with processing of complex shapes, results in deficits in remembering such shapes. Humans with focal lesions in the occipitotemporal cortical area, implicated in processing of specific visual features, also have difficulties remembering these features.

The activity of visual cortical neurons during working memory tasks also attests to their role in sensory maintenance. Such tasks similar to the delayed discrimination tasks described earlier, require the comparison of two stimuli separated by a memory delay. When such stimuli consist of moving random-dot stimuli and the monkeys are required to identify and remember the dot direction, the activity of MT neurons during the memory delay shows selectivity reflecting the remembered direction. This activity is pronounced early in memory delay but weakens toward its end and does not correlate with the animals' decision, suggesting that additional neural mechanisms are needed to account for the maintenance of motion signals. However, at the time of the comparison and decision, these neurons are strongly affected by the remembered direction, suggesting that throughout the task they have access to the remembered direction, and thus are likely to participate in the circuitry subserving storage. Similar behavior has been observed in inferotemporal (IT) cortex, associated with the processing of complex shapes, in that stimulus-selective activity occurs not only in response to specific shapes but also during the memory delay. During the comparison phase of the task, these neurons, like neurons in area MT, reflect the remembered shape. This type of activity is consistent with the possibility that visual cortical neurons actively participate the circuitry subserving storage of signals they process, although the nature of this participation awaits further study.

Functional imaging studies of visual working memory indicate that regions of the visual system are differentially recruited by working memory tasks, depending on the stimulus dimension that must be remembered. For example, working memory for the location versus the identity of visual stimuli, such as color patches and geometric shapes, recruits domain-specific memory-related activity in posterior cortical regions associated with the 'where' and 'what' visual streams, respectively. For location memory this includes dorsal occipital cortex, the intraparietal sulcus, and superior parietal lobule, whereas for object identity it includes ventral temporal and occipital cortex. Location memory tasks also typically recruit frontal regions associated with oculumotor control – the frontal eye fields and the supplementary eye fields.

Within the ventral streams, delay-period activity is segregated in a category-specific manner. For example, multiple images of faces and of naturalistic scenes can be presented as samples, with a postsample cue indicating which is the to-be-remembered category. When individuals are cued to remember the faces, delay-period activity is selectively elevated in the fusiform face area (FFA). When they are cued to remember scenes, delay-period activity is selectively elevated in the parahippocampal place area (PPA). Another study of working memory for faces featured three delay periods interposed between the presentation of the first and second, second and third, and third and fourth stimuli. The logic was that the multiple distracting events in this task might serve to 'weed out' from the first delay period activity that was not involved directly in storage, because only regions

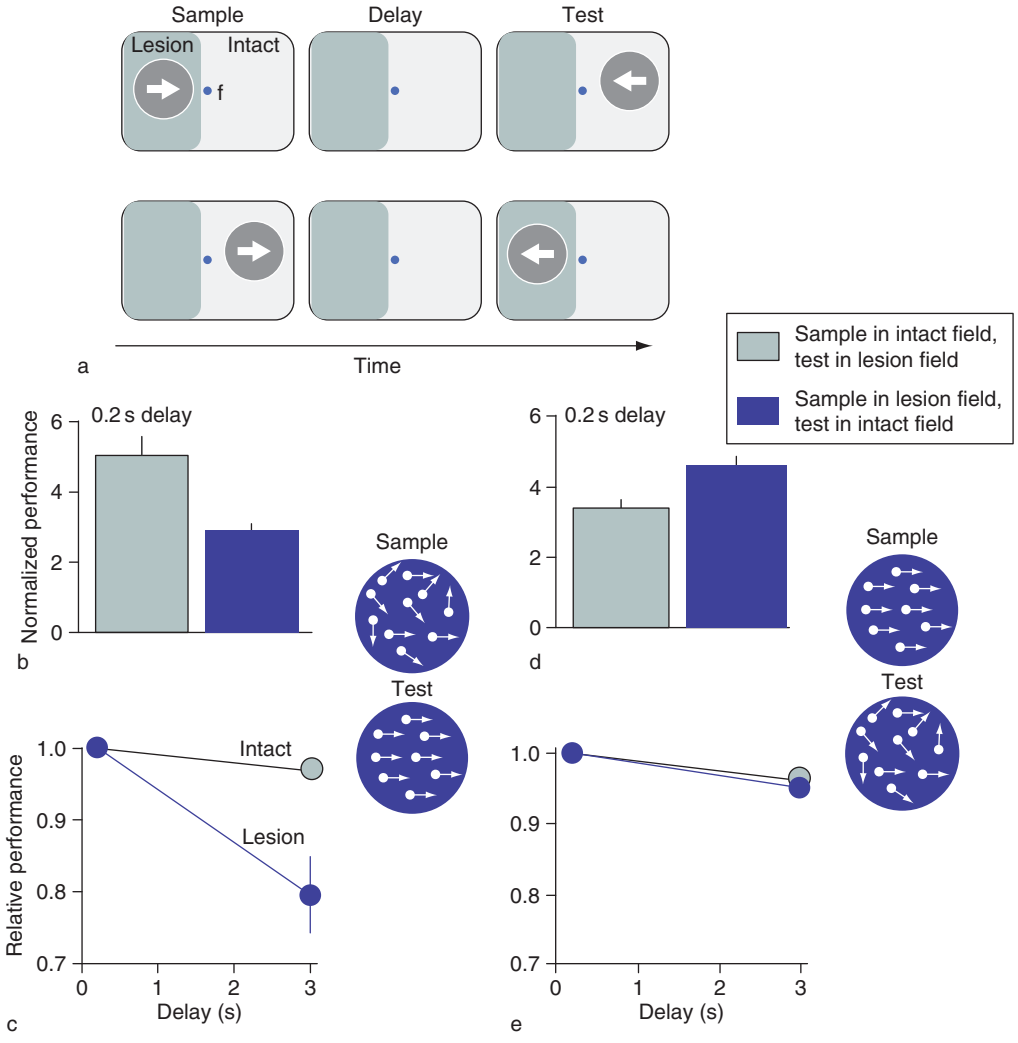

Figure 1 The role of cortical areas processing visual motion in remembering motion direction. (a) Monkeys with unilateral lesions of motion processing areas MT/MST performed a task requiring integration of complex motion and remembering its direction. In this task, sample and test stimuli were separated by a delay and positioned in opposite hemifields, so that one was placed in the lesioned and the other in the corresponding location in the intact hemifield. The monkeys reported whether sample and test moved in the same or in different directions by pressing one of two response buttons. During each trial the monkeys fixated a small target at the center of the display (blue dot) while attending to moving random dots (indicated by arrows) presented in the periphery. (b, d) Normalized direction range thresholds (1/(360-range threshold)/360) measured when the delay between sample and test was minimal (0.2 s). The sample or the test stimulus was composed of dots moving in a range of directions, while the other stimulus contained only coherent motion (stimulus configurations shown to the right of each set of plots). Performance was decreased whenever the stimulus containing noncoherent motion, and thus requiring integration, was placed in the lesioned field, demonstrating the importance of areas MT/MST for motion integration. (c, e) Effect of memory delay on performance for two direction range tasks. Performance was measured with both stimuli placed in the intact (gray circles) or in the lesioned hemifields (blue circles), either by varying the range of directions in the sample, while the test moved coherently (left plots), or by varying the range of directions in the test, while the sample moved coherently (right plots). Thresholds were normalized to the data measured at a 0.2 s delay. Error bars are SEM. A delay-specific deficit was present only when the remembered stimulus (sample) contained a broad range of directions and required integration. This result demonstrates that stimulus conditions requiring motion integration depend on intact areas MT/MST. However, coherently moving random dots can be discriminated and remembered at a normal level even in the absence of areas MT/MST. This shows that a cortical area involved in processing of sensory signals is also involved in their storage. Adapted from Bisley JW and Pasternak T (2000) The multiple roles of visual cortical areas MT/MST in remembering the direction of visual motion. *Cerebral Cortex* 10: 1053–1065.

with activity necessary for retaining the information to the end of the trial would maintain their activity across distracting stimuli. The results revealed that the posterior fusiform gyrus was the only region that retained the relevant signal during the last delay, immediately preceding the decision. Still other studies of working memory for faces have varied memory load (i.e., the number of items to be remembered),

on the logic that storage-related activity is sensitive to variations in load. These studies confirmed the importance of posterior fusiform gyrus in short-term storage.

Studies of working memory for spatial location have also addressed the mechanisms that support storage. One of these is attention-based rehearsal, a mechanism hypothesized to contribute to the short-term retention of locations via covert shifts of attention to the to-be-remembered location. These studies presented sample stimuli in one or the other visual hemifield while individuals fixated a central spot. Delay-period activity was greater in the hemisphere contralateral to the target location, an effect comparable to what one sees in studies of attention. This lateralized delay-period bias was strongest in extrastriate regions, decreased in magnitude across the parietal cortex, and was no longer reliable in frontal cortex. A second mechanism for the short-term retention of location information is prospective motor coding – the formulation, and then retention, of a motor plan for the acquisition of a target with, say, a saccade or a grasp. Electrophysiology studies in monkeys and neuroimaging studies in humans that encourage a prospective strategy localize this activity to frontal oculomotor regions, to prefrontal cortex, and to the caudate nucleus.

Tactile Working Memory

As with visual stimuli, tactile stimuli can be faithfully represented in working memory. Delayed discrimination of vibration stimuli can be performed with delays of many seconds, although the accuracy of this discrimination is maximal at short delays and decreases rapidly during the first 5 s of the delay. At longer delays, however, performance does not continue to deteriorate, suggesting that a two-stage memory process might be involved. Delayed discrimination of vibration stimuli can be disrupted by the application of transcranial magnetic stimulation (TMS) to human primary somatosensory cortex (area S1) during the initial portion of the delay period. This effect not only implicates area S1 in storage of vibration information, but also demonstrates the vulnerability of the storage mechanism early in the delay, an effect also observed in studies of working memory for visual motion.

In the awake behaving monkey, electrophysiological evidence for delay-period activity in S1 is mixed, with some laboratories reporting evidence in favor of it and others failing to find it. In area S2, however, there is no such equivocality, with strong evidence for stimulus frequency-specific activity persisting far into the delay period. Furthermore, responses of S2 neurons to the test stimulus contain information about the remembered stimulus, thereby reflecting the relationship between the two stimuli.

The ability to recognize and remember objects based on tactile input has also been studied with human neuroimaging. For example, positron emission tomography (PET) studies have revealed activation in the parietal operculum (area S2), an associative somatosensory area, during working memory tasks involving vibratory stimuli or palpated wire forms.

Auditory Working Memory

Although humans discriminate both pitch and loudness of sounds with high accuracy and remember these dimensions for many seconds, memory for pitch and loudness decline at different rates, suggesting that the two dimensions may be processed separately in auditory memory. The observed differences in the precision of memory for intensity and pitch parallel the findings in vision for contrast and spatial frequency (see earlier). As with vision, the use of interfering stimuli during the memory delay has revealed the nature of representation of auditory stimuli in memory. For example, memory for pitch can be disturbed by distractor tones, but only if these tones are within a narrow range of frequencies, relative to the frequency of the remembered stimulus, supporting the existence of separate pitch memory modules. Furthermore, sound frequency is likely to be stored separately from its location, suggesting that auditory memory obeys the same patterns as does auditory perception for physical parameters of the remembered stimuli.

In the auditory cortex, neuronal activity during the delay period of a delayed two-tone discrimination carries information about the frequency of the sample tone. In addition, the response to the second tone depends on whether its frequency matches that of the first, implicating neurons in auditory cortex in the circuitry subserving working memory for tone frequency. Participation of auditory cortex in auditory working memory is also supported by studies of the human using magnetoencephalography (MEG).

Memory-related activity has also been seen in associative auditory cortex of humans during tasks requiring memory for differences in pitch and tonal sequences; the right superior temporal lobe was associated with tonal sequences, and pitch judgments were associated with increased right frontal lobe activation. Interestingly, the auditory cortex appears to be maximally active early in the delay, whereas regions in the supramarginal gyrus and parts of the cerebellum are activated later, suggesting that for audition, too, working memory may be accomplished in multiple stages.

The involvement of regions processing auditory information in working memory is also supported by selective deficits in patients with damage to auditory associative cortex in the right temporal lobe. These patients are impaired primarily on tone discrimination with long, distracted delays.

Multiple Encoding in Working Memory

It seems unlikely that the representational bases of working memory are limited to the domain in which stimuli are perceived. In addition to the multiple serially ordered stages in working memory for visual, tactile, and auditory domains, sensory information can also be represented in parallel in multiple formats. Thus, as reviewed earlier, working memory for location can be accomplished via retrospective sensory memory-based mechanism, as well as a prospective motor planning mechanism. Analogously, there is growing evidence that working memory for visual identity recruits verbal (and, perhaps, semantic) representations in addition to the ventral stream visual mechanisms already discussed.

Working Memory and Prefrontal Cortex

For many years PFC has been strongly associated with the storage of sensory information in working memory. Consistent with this idea are the facts that PFC is directly interconnected with cortical areas processing fundamental sensory dimensions, and that PFC neurons often display stimulus-related activity during the memory delay. However, a growing body of evidence suggests that PFC is more likely to play a key role in directing attention to behaviorally relevant sensory signals and in making decisions concerning these signals, than in directly supporting their retention in short-term and working memory. Consistent with this view is the fact that extensive damage to prefrontal cortex does not eliminate the ability to perform simple short-term memory tasks, such as the digit span and spatial span variants of immediate serial recall. The limited effect of such lesions is apparent on delay tasks if during the memory period there is no interference and no requirement to keep track of multiple items. Also consistent with this view is the finding that although the delay activity of PFC neurons recorded in monkeys performing a delayed direction discrimination task (see earlier) was direction-specific, it did not predict decisions about stimulus direction that they made at the end of the trial. The contribution of these neurons to the task became apparent only after the end of the memory delay, when the monkeys compared test direction with that of the remembered sample. Furthermore, direction selectivity in PFC neurons, prevalent during the task, was greatly reduced when the monkeys were not required to use the information about stimulus direction. This work showed that PFC neurons are capable of gating sensory signals according to their behavioral relevance, it but does not support the key role for PFC neurons in the maintenance of these signals. Other studies, designed to isolate the contributions of attention and response selection from those of sensory storage, indicate a stronger role for the former in PFC.

In the human neuroimaging literature there is some debate about stimulus domain- or category-specific segregation of delay-period activity in PFC – for example, comparable patterns of delay-period activity during tasks involving memory for location versus memory for objects or memory for faces versus memory for scenes. The same is true for tasks that involve haptic object encoding versus those that involve visual object encoding. Similarly, in tasks requiring cross-modal integration, individual PFC neurons can be active during the memory delay following the presentation of different stimulus modalities. These examples suggest that PFC assists early-level sensory processing regions to form supramodal mental representations of objects and/or of task contingencies when the relevant stimuli belong to more than one sensory domain.

Another strategy for evaluating the neural basis of sensory storage is to require the retention of information across multiple delay periods interrupted with intervening distracting stimuli. This approach has revealed that neurons in the monkey PFC, but not IT cortex, could sustain a representation of the sample stimulus across multiple delays. Robust sample-specific activity across multiple delay periods is not limited to PFC, however, and has also been observed in other regions, including the temporal pole and the entorhinal cortex of the medial temporal lobe. Functional magnetic resonance imaging (fMRI) studies in humans remembering faces showed that only posterior fusiform gyrus retains sustained activity across all three delay periods, although many regions displayed activity during the first delay period. An analogous result was seen in the individuals remembering spatial locations, with distraction-spanning delay-period activity distributed across multiple areas, including intraparietal sulcus, superior parietal lobule, frontal eye fields, and supplementary eye fields in frontal cortex (Figure 2). Finally, repetitive TMS (rTMS) applied to PFC has failed to disrupt working memory performance, but has altered performance when applied to parietal cortex.

Although the weight of evidence is inconsistent with a storage role for PFC, virtually every electro-

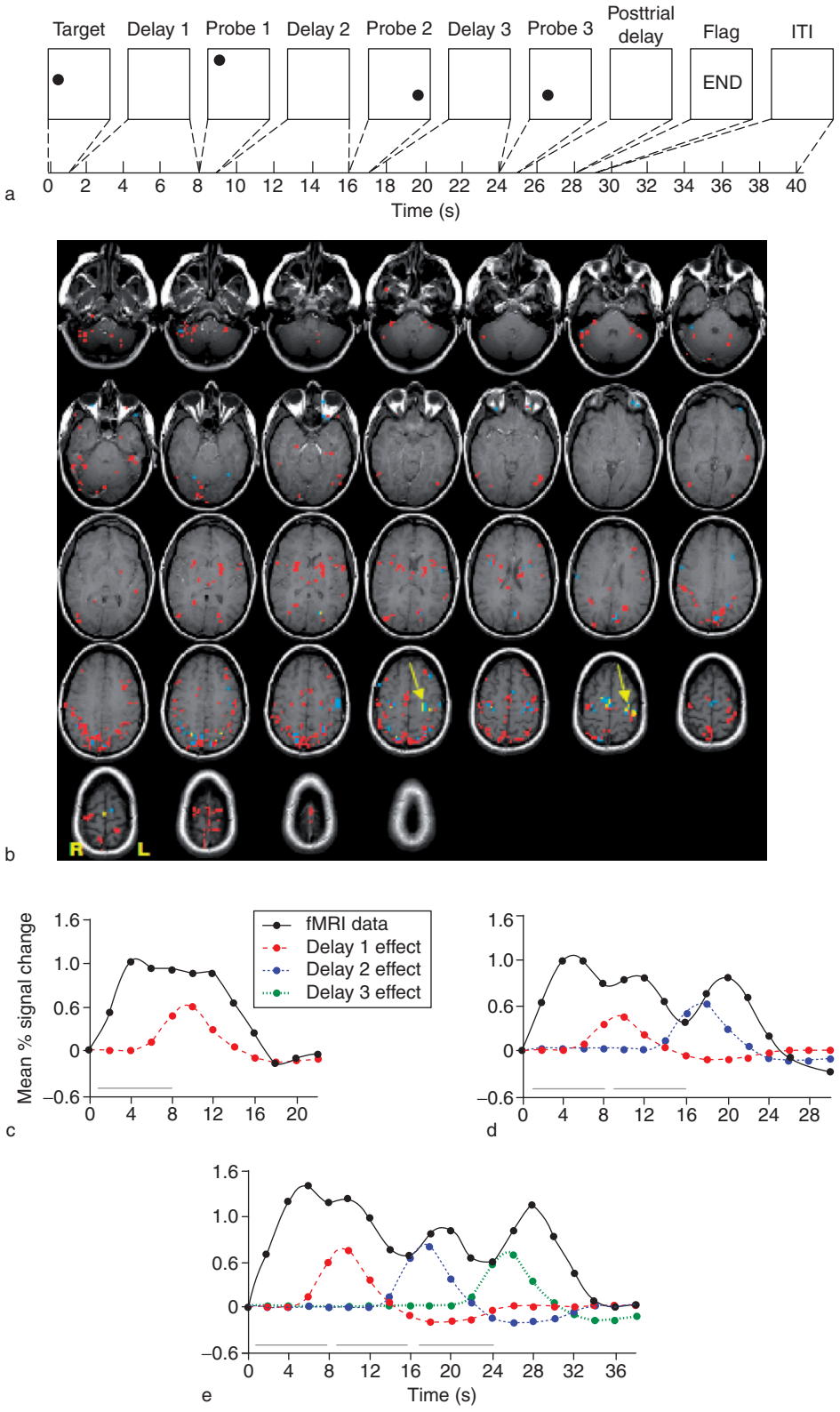

Figure 2 A functional magnetic resonance imaging experiment demonstrating sustained memory-related activity for location across multiple delay periods. (a) Behavioral task. Individuals view and encode the target location, then, after a 7 s delay, indicate with a button press whether the probe does or does not appear in the same location. One-third of the trials end after probe 1. On two- and three-delay trials, the offset of probe 1 is followed by another delay period, after which individuals evaluate the location of probe 2 with respect to the target. On one- and two-delay trials, the 'END' message appears at times 12 and 20 s, respectively. (b) Results from a single representative individual. Voxels in red showed sustained delay-period activity for delay 1 only. Voxels in blue are the subset of voxels

physiological and neuroimaging study of primate working memory finds delay-period activity in this region of the brain. Thus, it is important to address the role of this activity. If this activity does not represent the storage of information, what are alternative explanations of its function? One possibility is that PFC plays an important role in the control of potentially disrupting effects of distraction and interference during working memory tasks. Lesion, electrophysiology, and fMRI studies have all provided evidence that dorsolateral PFC can accomplish this by controlling the gain of activity in sensory processing regions, such that the delay-period processing of potentially distracting stimuli is suppressed. Analogously, a region of ventrolateral PFC has been implicated in the control of proactive interference, the deleterious effect of previous mental activity on current task performance. In addition, an important function of delay activity in PFC may be the attentional selection of task-relevant information, as well as in planning for the response. This is consistent with the well-characterized role for PFC in the biasing of stimulus–response circuits so that novel or less salient behaviors can be favored over well-learned associations and behavioral routines. Finally, the potential role of PFC in the integration of cognitive and motivational factors is supported by the modulation of delay-period activity of its neurons during spatial memory tasks by the type of the anticipated reward.

Conclusion

The emerging picture of the neural basis of working memory is of a class of behaviors that does not depend on one or more functionally specialized regions. Rather, working memory is supported by the coordinated activity of circuits responsible for the sensory processing of the critical information, and, to varying degrees, those that control the flexible allocation of attention and the selection of task-relevant behavior.

See also: Attentional Functions in Learning and Memory; Cognition: An Overview of Neuroimaging Techniques; Event-Related Potentials (ERPs); Hippocampus and Neural Representations; Prefrontal Cortex: Structure and Anatomy; Somatosensory Perception; Synaptic Plasticity and Place Cell Formation; Working Memory: Capacity Limitations.

Further Reading

Baddeley A (1992) Working memory. *Science* 255: 556–559.

Bisley JW and Pasternak T (2000) The multiple roles of visual cortical areas MT/MST in remembering the direction of visual motion. *Cerebral Cortex* 10: 1053–1065.

Cowan N (2005) *Working Memory Capacity*. Hove, UK: Psychology Press.

Fuster J (1995) *Memory in the Cerebral Cortex*. Cambridge, MA: MIT Press.

Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14: 477–485.

Hebb D (1949) *Organization of Behavior*. New York: Wiley.

Miller EK and Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24: 167–202.

Pasternak T and Greenlee M (2005) Working memory in primate sensory systems. *Nature Reviews Neuroscience* 6: 97–107.

Petrides M (2000) Dissociable roles of mid-dorsolateral prefrontal and anterior inferotemporal cortex in visual working memory. *Journal of Neuroscience* 20: 7496–7503.

Postle BR (2006) Distraction-spanning sustained activity during delayed recognition of locations. *NeuroImage* 30: 950–962.

Postle BR (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139: 23–38.

Postle BR and Hamidi M (in press) Nonvisual codes and nonvisual brain areas support visual working memory. *Cerebral Cortex-Advance Access*. published online on December 5, 2006.

Romo R and Salinas E (2003) Flutter discrimination: Neural codes, perception, memory and decision making. *Nature Reviews Neuroscience* 4: 203–218.

Todd J and Marois R (2005) Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cognitive, Affective & Behavioral Neuroscience* 5(2): 144–155.

Zaksas D and Pasternak T (2006) Directional signals in the prefrontal cortex and area MT during a working memory for visual motion task. *Journal of Neuroscience* 26: 11726–11742.

with delay-period activity sustained across delay 1 and delay 2. Voxels in yellow are the still smaller subset with delay-period activity sustained across all three delays. Note that although many regions, including the prefrontal cortex, show delay-period activity during delay 1, only dorsal-stream parietal and frontal oculomotor regions sustain this delay-period activity across all three delays. Arrows highlight the voxels from left frontal eye field (the activity of which is shown in panels c, d, and e). (c) Activity from three-delay voxels in the left frontal eye field, averaged across one-delay trials. 'Delay 1 effect' reflects the estimated magnitude of delay 1 activity. Gray bar along the horizontal axis indicates the duration of the delay period. (d) Activity from these same frontal eye field voxels averaged across two-delay trials. Graphical conventions are the same as in panel b. (e) Activity from these same frontal eye field voxels averaged across three-delay trials. Graphical conventions are the same as in panel c. Adapted from Postle BR (2006) Distraction-spanning sustained activity during delayed recognition of locations. *NeuroImage* 30: 950–962.