Neural evidence for a 3-state model of visual short-term memory

Derek Evan Nee [a,*], John Jonides [b]

[a] Helen Wills Neuroscience Institute, University of California, Berkeley, USA
[b] Department of Psychology, University of Michigan, USA

ARTICLE INFO

ABSTRACT

Recent research has suggested that short-term memory (STM) can be partitioned into three distinct states. By this model, a single item is held in the focus of attention making it available for immediate processing (focus of attention), a capacity-limited set of additional items is actively maintained for future processing (direct access region), and other recently presented information is passively active, but can nevertheless influence ongoing cognition (activated portion of long-term memory). While there is both behavioral and neural support for this 3-state model in verbal STM, it is unclear whether the model generalizes to non-verbal STM. Here, we tested a 3-state model of visual STM using fMRI. We found a triple dissociation of regions involved in the access of each hypothesized state. The inferior parietal cortex mediated access to the focus of attention, the medial temporal lobe (MTL) including the hippocampus mediated access to the direct access region, and the left ventrolateral prefrontal cortex (VLPFC) mediated access to the activated portion of long-term memory. Direct comparison with previously collected verbal STM data revealed overlapping neural activations involved in the access of each state across different forms of content suggesting that mechanisms of access are domain general. These data support a 3-state model of STM.

Introduction

At the core of complex cognition is the ability to hold information in mind for brief periods of time so that it can be processed in the absence of external stimulation. This ability is referred to as short-term memory (STM). The centrality of STM is evidenced by the strong relationship between variations in STM-capacity and higher-order cognitive skills such as reasoning, language comprehension, and intelligence (Carpenter et al., 1990; Daneman and Carpenter, 1980; Daneman and Merikle, 1996). Furthermore, cognitive training that increases the capacity of STM also increases fluid intelligence (Jaeggi et al., 2008). Hence, understanding STM has far-reaching implications, and an appropriate theoretical account of STM is central for any model of complex cognition.

A number of models have attempted to provide a theoretical description of STM. Arguably, the most influential such model is that of Baddeley (Baddeley, 1986, 2003; Baddeley and Hitch, 1974). This model assumes separable STM buffers for different kinds of content (e.g. verbal and visual). These buffers are presumed to form distinct stores that can be distinguished from long-term memory (LTM). Collectively, the various STM buffers and the executive processes that act upon the buffers are referred to as working memory. Since the turn of the millennium, separate *store* distinctions between STM and

LTM have grown tenuous with most recent models assuming that STM consists of activated LTM representations (see Jonides et al., 2005, 2008; Nee et al., 2008; Postle, 2006 for reviews). By these accounts, the same neural tissue is responsible for representing both STM and LTM with the activity-level of the neural tissue distinguishing different *states* that correspond to STM and LTM. Thus, different forms of memory reflect different states rather than stores. Various authors have posited such 2-state models of memory (Cowan, 2001; McElree, 2006; McElree and Dosher, 1989), with different hypotheses regarding the number of representations that can be actively maintained. According to the model of Cowan (2001), approximately 4 ± 1 representations can be actively maintained at a given time while the model of McElree (2006) posits that only a single representation can be actively maintained. Behavioral data have demonstrated sharp performance discontinuities that provide support for both accounts (see Jonides et al., 2008 for a review) suggesting that each proposal is at least partially correct. As a result, Oberauer (2002) has suggested that both single-item and 4 ± 1 capacities exist and that each of these capacities reflects different states in STM.

According to the model of Oberauer (2002), STM consists of three different states. First, a single item resides in the focus of attention (FA) that is available for immediate cognitive processing. The FA has access to a capacity-limited number of items that are actively maintained for use in ongoing cognition referred to as the direct access region (DAR). The FA can be switched flexibly among the items in the DAR in order to highlight a single item for further processing. Finally, the DAR can be distinguished from the activated

portion of LTM (aLTM) consisting of information that is passively maintained. That is, representations in the aLTM retain some level of activation[1] due to recent presentation or due to associations with information in the FA and DAR. Critically, however, representations in the aLTM lack the top-down biasing that maintains items in the FA and the DAR. The influence of representations in the aLTM on cognition is evidenced by phenomena such as priming and proactive interference. In support of this 3-state model, Oberauer (2002) demonstrated behavioral costs associated with each hypothesized memory state. This model then provides a holistic synthesis of different accounts of STM while also demonstrating the continuum among attention, STM, and LTM.

Recently, we and others have documented neural evidence for Oberauer's 3-state model of memory (Nee and Jonides, 2008, 2011; Oztekin et al., 2009, 2010). These studies have employed serial item-recognition tasks that use temporal recency as a mechanism to distinguish different qualitative states of memory. For example, we presented subjects with a 6-word list followed immediately by a recognition probe (Nee and Jonides, 2011). Rapid presentation of targets and a brief retention interval minimized strategic processing that may have produced chunking or disrupted the temporal order via rehearsal. Hence, we assumed that the most recently presented item was retained in the FA. We assumed further that the next most recent items up to a capacity-limit were retained in the DAR with supra-capacity items residing in the aLTM. Behavioral recognition data demonstrated a sharp performance discontinuity indicating a clear distinction between the putative DAR and the aLTM within hypothesized ranges. By examining fMRI responses to the recognition probes, we could query the neural mechanisms involved in accessing information in the three hypothesized states. We reasoned that if the 3-state model was correct, a triple dissociation should demonstrate qualitatively distinct regions mediating access to each of the three states. This is exactly what we found. Accessing the FA involved inferior parietal and inferior temporal activations, accessing the DAR involved the medial temporal lobe (MTL), and accessing aLTM involved the ventrolateral prefrontal cortex (VLPFC). Similar patterns have been documented in related studies (Nee and Jonides, 2008; Oztekin et al., 2009, 2010). Hence, these data provide neural support for a 3-state model of memory.

An important limitation of previous research is that evidence for a 3-state model of memory has relied entirely on studies employing verbal material. As a result, it is unclear whether the 3-state model of memory is a domain-general principle or whether it applies solely to the verbal domain. For instance, while a capacity limit for verbal STM is evident at approximately 4 ± 1 items when rehearsal is prevented, this number increases to the well-known 7 ± 2 when rehearsal is permitted (Cowan, 2001; Miller, 1956). Although previous fMRI studies investigating different memory states have attempted to minimize rehearsal strategies, it is nevertheless possible that phonological coding may have imparted a unique status to the verbal information. If so, this could account for a third state that may not exist in other forms of STM. Moreover, behavioral data using serial recognition paradigms tapping visual STM have often demonstrated particularly sharp performance drop-offs after just one or two items with plateaus thereafter (Broadbent and Broadbent, 1981; Hanna and Loftus, 1993; Hay et al., 2007; Kerr et al., 1999; Phillips and Christie, 1977). Such data appear consistent with 2-state models that assume distinctions only between a single-item FA and LTM (McElree, 2006; McElree and Dosher, 1989). Furthermore, recent meta-analyses have revealed domain specificity of both storage (Rottschy et al., 2012) and executive control of STM (Nee et al., 2013) demonstrating that different neural correlates and potentially distinct mechanisms underlie STM of different content. As

a result, it is critical to examine putative 3-state distinctions outside of the verbal domain to investigate whether the model generalizes.

The present study was designed to test a 3-state model of visual STM. To do so, we adapted our paradigm that established a 3-state model of verbal STM (Nee and Jonides, 2011) to test visual STM. On each trial, subjects were sequentially presented with five faces followed shortly thereafter by a recognition probe. In accordance with previous research (Nee and Jonides, 2008, 2011; Oztekin et al., 2009, 2010), we assumed that probes matching the most recently presented item queried access to the FA. Because we expected visual STM to be more variable than verbal STM, we separately estimated each individual's capacity and assumed that probes querying items within capacity-limits provided an assay of access to the DAR. Finally, we assumed that probes matching supra-capacity items measured access to the aLTM. If a 3-state model applies to visual STM, we would expect a triple dissociation of neural regions responsive to the access of each memory state. If, instead, a 2-state model applies to visual STM, we would expect to find only a double dissociation. If access is domain general across different forms of STM, we would expect the same regions involved in accessing different visual memory states as we found with verbal memory states (Nee and Jonides, 2011). That is, we would expect inferior parietal areas to correlate with accessing the FA, the MTL to correlate with accessing the DAR, and the VLPFC to correlate with accessing the aLTM. If access is domain specific, we would expect different areas involved in accessing visual STM as compared to verbal STM. Hence, our design permitted the examination of a 3-state model of visual STM while comparisons to our previously collected verbal data enabled the investigation of domain generality/specificity of the 3-state model.

Materials and methods

Participants

Data were collected from 25 right-handed adults (13 females; ages 18–25; mean age = 19.9 years old). Due to excessive motion, data from 1 subject was excluded from fMRI analysis. Informed consent was obtained for all subjects in accordance with the Institutional Review Board at the University of Michigan. Subjects received $20/h as well as a bonus for fast and accurate performance.

Materials and procedure

The task is depicted in Fig. 1. Subjects performed a serial item-recognition task modeled after paradigms that have been used to examine different states of verbal STM (McElree, 2006; McElree and Dosher, 1989; Nee and Jonides, 2008, 2011; Oztekin et al., 2009, 2010). On each trial, subjects were sequentially presented five faces (target set), followed by a brief mask and a recognition probe. Subjects responded with a keypress regarding whether the recognition probe was a member of the target set or not. Half of the probes matched a member of the target set (match probe) and half did not (non-match probe). Match probes were equally distributed among the five serial positions (SPs). Faces were drawn randomly without replacement from a set of 179 gray-scale images (see Fig. 1 for examples). When exhausted, the set was re-randomized. A separate set of 94 gray-scale faces was used for practice runs.

Each target was presented for 1000 ms with a 250 ms inter-stimulus interval spaced between successive targets. Each target was presented in a separate location drawn from a set of six potential locations arranged hexagonally around the fixation cross. Each location was placed equidistant from the fixation cross and equidistant from each adjacent location. The locations were irrelevant for the task and were used to alleviate visual overwriting that might make the task too difficult for subjects and mask potential state differences. Pilot data and previous research were consistent

---

[1] Here, the term "activation" is used as a theoretical construct to denote the accessibility of a representation. It does not necessarily reflect active neural firing and instead may be realized by synaptic potentiation.
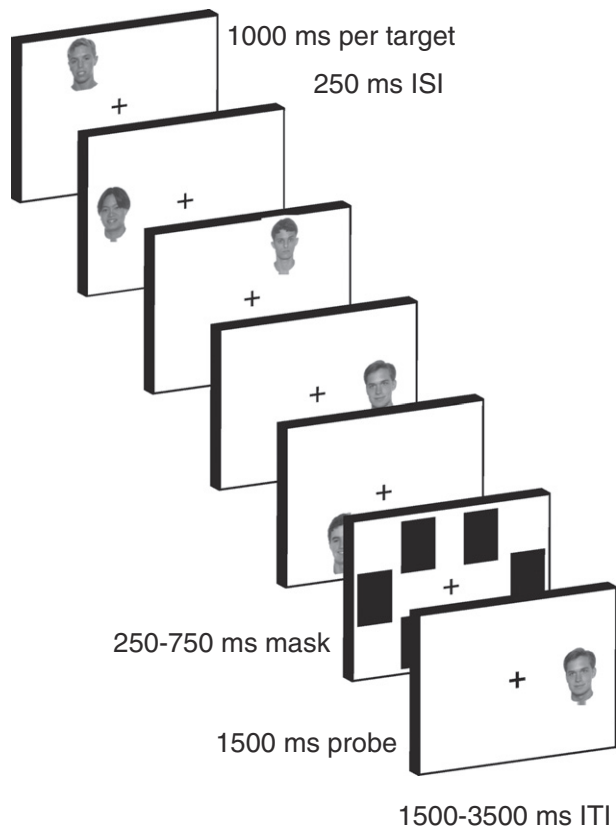
1000 ms per target

250 ms ISI

250-750 ms mask

1500 ms probe

1500-3500 ms ITI

Fig. 1. Task depiction. On each trial, subjects were sequentially presented five gray-scale target faces. Targets were presented in one of six locations arranged hexagonally around fixation. After a brief mask, a recognition probe appeared to which subjects made a match/non-match decision. Variations in the serial position (SP) of the probe indexed access to different putative states of memory. Probes matching the most recently presented target (SP −1) were presumed to measure access to the focus of attention. Subsequently distant probes up to capacity-limits were presumed to measure access to the direct access region (e.g. SPs −2 and −3). Probes distant to these were presumed to measure access to the activated portion of long-term memory (e.g. SPs −4 and −5).

with this idea (Broadbent and Broadbent, 1981). After presentation of the target set, a mask appeared for 250 to 750 ms consisting of black rectangles over each of the six potential target locations. The mask was immediately followed by a recognition probe presented in one of the six locations and lasting 1500 ms. Match probes always appeared in the location in which the corresponding target had previously appeared. Non-match probes appeared in a random target location, but with a stimulus that had not appeared for at least 3 trials in order to minimize proactive interference. The probe was followed by a 1500 to 3500 ms inter-trial interval. During the last 500 ms of the inter-trial interval, the fixation cross changed to red to alert the subject that the next trial was about to begin. A black fixation cross remained on the screen during all other times and subjects were instructed to gaze at the fixation cross whenever they were not encoding a target or probe. Subjects were informed that this was especially important during the mask interval. This procedure ensured control for potentially confounding effects of differential saccades during the probe interval. Eye movements were recorded to monitor adherence to instructions. Subjects performed 6 runs of 30 trials each. Immediately prior to entering the scanner, subjects performed 2 runs of 30 trials each as practice to familiarize them with the task.

In accordance with previous research, we assumed that different putative states of STM could be examined by varying the SP of the match probe. In keeping with prior work (McElree, 2006; McElree and Dosher, 1989; Nee and Jonides, 2008, 2011; Oztekin et al., 2009,

2010), we assumed that probes matching the most recently presented item (SP −1) assessed access to the FA. Previously, we had assumed that probes matching SPs −2 and −3 reflected access to the DAR based upon clear behavioral patterns (Nee and Jonides, 2011). However, pilot work demonstrated that the present paradigm using face stimuli produced more variable behavioral performance than our previous paradigm that used words. As a result, we tailored putative assessments of the DAR based upon individual differences in capacity. For each subject, we estimated Cowan's K (Cowan, 2001) in the serial item-recognition task (K_SIRT) as a measure of capacity through the formula (correct hit rate + correct rejection rate − 1) × N where N corresponds to the number of items (i.e. 5 in the present study). K_SIRT ranged from 1.72 to 4.39 items among subjects with an average of 2.76. Consistent with Oberauer (2002), we reasoned that K_SIRT reflects the capacity of the DAR and that supra-capacity items would reside in the aLTM. Hence, we rounded K_SIRT estimates and assigned SPs −2 to −K_SIRT to reflect retrieval from the DAR, and SPs −K_SIRT+1 to −5 to reflect retrieval from aLTM. For example, in a subject with a K_SIRT of 3, SPs −2 and −3 were assumed to assess the DAR and SPs −4 and −5 were assumed to assess the aLTM. Similar, but more variable neural results were obtained using comparable analyses that assumed a capacity of 3 for all subjects.

After scanning, subjects performed a standard change detection task (CDT) to provide an independent estimate of STM capacity (Luck and Vogel, 1997). In the change detection task, subjects were presented with an array of colored squares of varying set sizes. Following a brief retention interval, the array was re-presented with one square encircled. The subject's task was to indicate whether the color of the encircled square changed from sample to test. Set sizes varied from two to eight items. Capacity was measured using Cowan's K formula described above for each set size and the maximum K value was assumed to be the subject's capacity (K_CDT).

Image acquisition and preprocessing

Images were acquired on a GE Signa 3 T scanner equipped with a 4-channel head coil. Head movement was minimized using foam padding and a cloth restraint strapped across participants' foreheads. Experimental tasks were presented using E-Prime software version 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA). Eye tracking was performed using ViewPoint (Arrington Research, Inc., Scottsdale, AZ). Due to technical issues, eye tracking could not be performed on 2 subjects.

Functional T2*-weighted images were acquired using a spiral sequence with 43 contiguous slices with $3.44 \times 3.44 \times 3$ mm voxels (repetition time, or TR = 2000 ms; echo time, or TE = 30 ms; flip angle = 90°; field of view, or FOV = 220 mm^2). A T1-weighted gradient-echo anatomical overlay was acquired using the same FOV and slices (TR = 250 ms, TE = 5.7 ms, flip angle = 90°) to improve co-registration between the high-resolution anatomical image and functional images. Additionally, a 124-slice high-resolution T1-weighted anatomical image was collected using spoiled-gradient-recalled acquisition (SPGR) in steady-state imaging (TR = 9 ms, TE = 1.8 ms, flip angle = 15°, FOV = 250–260 mm^2, slice thickness = 1.2 mm).

Functional data were spike-corrected to reduce the impact of artifacts using AFNI's 3dDespike (http://afni.nimh.nih.gov/afni). Subsequent processing and analyses were done using SPM5 (http://www.fil.ion.ucl.ac.uk/spm/). Functional images were corrected for differences in slice timing using sinc-interpolation and head movement using a least-squares approach and a 6-parameter rigid body spatial transformation. Structural data were co-registered to the functional data and segmented into gray and white-matter probability maps (Ashburner and Friston, 1997). These segmented images were used to calculate spatial normalization parameters to the MNI template; these were subsequently applied to the functional data. As part of

spatial normalization, the data were resampled to $2 \times 2 \times 2 \text{ mm}^3$. 8-mm full-width/half-maximum isotropic Gaussian smoothing was applied to all functional images prior to analysis using SPM5. All analyses included a temporal high-pass filter (128 s), correction for temporal autocorrelation using an autoregressive AR(1) model, and each image was scaled to have a global mean intensity of 100.

Imaging analysis

fMRI data were analyzed using the general linear model implemented in SPM5. Regressors of main interest were locked to the onset of recognition probes and convolved with a canonical hemodynamic response function. Separate regressors were included for probes matching each of the five serial positions. Similarly, separate regressors were included for non-match probes with serial position indicated by the location of the non-match probe (i.e. non-match SP −2 would indicate a non-match probe presented in the same location as the SP −2 target). Error trials were modeled separately and excluded from foregoing analyses. Epoch regressors covering the encoding period were included to capture variance associated with stimulus encoding. For subjects demonstrating greater than 3 mm of motion across a session or greater than 0.5 mm of motion between TRs, 24 motion regressors were included to capture linear, quadratic, differential, and squared differential residual motion variance (Lund et al., 2005).

Parameter estimates for match probes were grouped into putative memory states as described above and submitted to a second-level one-way ANOVA. Within the second-level model, we performed three separate contrasts to examine a potential triple dissociation that would support a 3-state model of memory. These contrasts looked for brain areas more active as a function of retrieval from one memory state relative to other memory states. To ensure that activations were not confounded with overall difficulty, we constrained voxels-of-interest to show no difference between states of non-interest. Concretely, we looked for FA-specific voxels through the contrast FA > (DAR + aLTM)/2 (p < 0.05 family-wise error (FWE) corrected) with the added constraint of no difference between DAR and aLTM (DAR = aLTM; i.e. voxels showing DAR−aLTM at p > 0.05 uncorrected, two-tailed). Similarly, DAR-specific voxels were explored through DAR > (FA + aLTM)/2 & FA = aLTM, and aLTM-specific voxels were explored through aLTM > (FA + DAR)/2 & FA = DAR. For whole-brain analyses, data were thresholded at p < 0.001 at the voxel level with a 75 voxel extent producing p < 0.05 FWE corrected according to simulation using AlphaSim.

Previous studies have indicated that the MTL is an important region-of-interest. To explore activations in the MTL that may not survive our strict whole-brain threshold, we performed separate searches within the MTL. The MTL was defined as the bilateral hippocampi and parahippocampal gyri as specified by the AAL atlas implemented in SPM5 through WFU Pick Atlas (Maldjian et al., 2003). Within the MTL, we repeated the above analyses at a reduced voxel-wise threshold (p < 0.05) with an extent threshold of 190 voxels producing p < 0.05 FWE correction according to simulations using AlphaSim.

To examine the domain generality of areas involved in accessing different memory states, we performed conjunction analysis (Nichols et al., 2005). The above contrasts were repeated on our previously collected data using word stimuli (Nee and Jonides, 2011). We then looked for areas significantly active for the same contrast across verbal STM and visual STM. Areas of overlap greater than 20 voxels in extent are reported.

Results

Behavioral results

Behavioral data were analyzed to investigate hypothesized differences between retrieval from different memory states indexed by differences in SP. We performed 2-way ANOVAs with factors of SP (−1 to −5) and probe-type (match, non-match). Reaction time (RT) data were analyzed on correct trials only. As expected, performance declined with SP evidenced by a main effect of SP in both accuracy (F(4,96) = 49.42, p < 0.0001) and RT (F(4,96) = 119.9, p < 0.0001). Performance was also significantly better on non-match relative to match trials in accuracy (F(1,24) = 13.52, p < 0.005) with a non-significant trend in RT (F(1,24) = 3.38, p < 0.08). Finally, a SP × probe-type interaction was present in both accuracy (F(4,96) = 13.59, p < 0.0001) and RT (F(4,96) = 10.68, p < 0.0001).

To unpack the SP × probe-type interactions, we examined the data in more detail. The RT data demonstrated three distinct patterns (Fig. 2A). At SP −1, match decisions were made faster than non-match decisions (t(24) = 3.13, p < 0.005). However, no RT differences were present at SPs −2 and −3 (both t(24) < 0.65, p > 0.5). Finally, at SPs −4 and −5, non-match decisions were made more quickly than match decisions (both t(24) > 3.5, p < 0.005). Interestingly, these three distinct patterns mapped directly to the three putative states of memory. The same pattern was confirmed when the data were grouped according to individually measured capacities (Fig. 2B). While these distinctions were not necessarily predicted by the 3-state model, they nevertheless indicate the possibility of three qualitatively distinct states.

In accuracy, performance on non-match probes declined from SP −1 to SP −2 (t(24) = 4.24, p < 0.0005), but remained stable thereafter (all other pairwise t(24) < 1.5, p > 0.14). By contrast, performance continued to decline on match trials through SP −4 (SP −1 to −4, all step-wise t(24) > 4.1, p < 0.0005 before reaching a plateau (SP −4 vs −5, t(24) = 0.5, p > 0.6). Notably, match performance on SPs −4 and −5 did not differ from chance (both t(24) < 1.5, p > 0.14). However, chance performance does not indicate a lack of memory for SP −4 and −5 items. If this was the case, the hit rate for probes matching SPs −4 and −5 should be no different than the false alarm rate for SP −4 and −5 non-match probes (i.e. d' = 0). That is, match probes should appear subjectively no different than non-match probes. Instead, these differed a great deal (both t(24) > 8.3, p < 0.0001). Furthermore, that non-match performance was constant across SP −2 to −5 precludes the possibility that subjects simply adopted a "no" bias to enhance non-match performance for SPs −4 and −5. Together, these data indicate that a memory trace remained even on SPs −4 and −5 (Fig. 2C). When grouped by individually measured capacities, the data indicated no performance difference between match and non-match probes for accessing the FA (t(24) < 0.25, p > 0.8) or the DA (t(24) = 1.02, p > 0.3), but a significant difference for accessing the aLTM (t(24) = 5.06, p < 0.0005). These data indicate a dramatic performance drop-off when accessing the aLTM relative to the FA and the DA.

Next, we correlated performance on the serial item-recognition task with capacity-estimates derived from the CDT. Cowan's K on the serial item-recognition task (K_SIRT) averaged 2.76 with a range of 1.72 to 4.39 indicating that subjects could accurately maintain two to four faces in STM. Capacity measured from the CDT (K_CDT) correlated positively with K_SIRT (Spearman's ρ = 0.45, p < 0.05) demonstrating a common construct measured by both tasks (Supplemental Fig. 1A). To determine whether this relationship held for all putative memory states, we correlated K_CDT with d' computed separately for each state (FA, DAR, aLTM). When grouped on the basis of average K_SIRT (i.e. FA = SP −1, DAR = SPs −2 & −3, aLTM = SPs −4 & −5), K_CDT correlated with retrieval from DAR (Spearman's ρ = 0.43, p < 0.05; Supplemental Fig. 1C) and aLTM (Spearman's ρ = 0.40, p < 0.05; Supplemental Fig. 1D), but did not correlate with retrieval from FA (Spearman's ρ = 0.06, p > 0.75; Supplemental Fig. 1B). However, these relationships were no longer significant when grouping on the basis of each individual's K_SIRT (both ρ < 0.32, p > 0.1). This result suggests that the tailored groupings successfully minimized capacity-variability that could complicate imaging analyses. Thus, we utilize tailored groupings for subsequent imaging analyses.
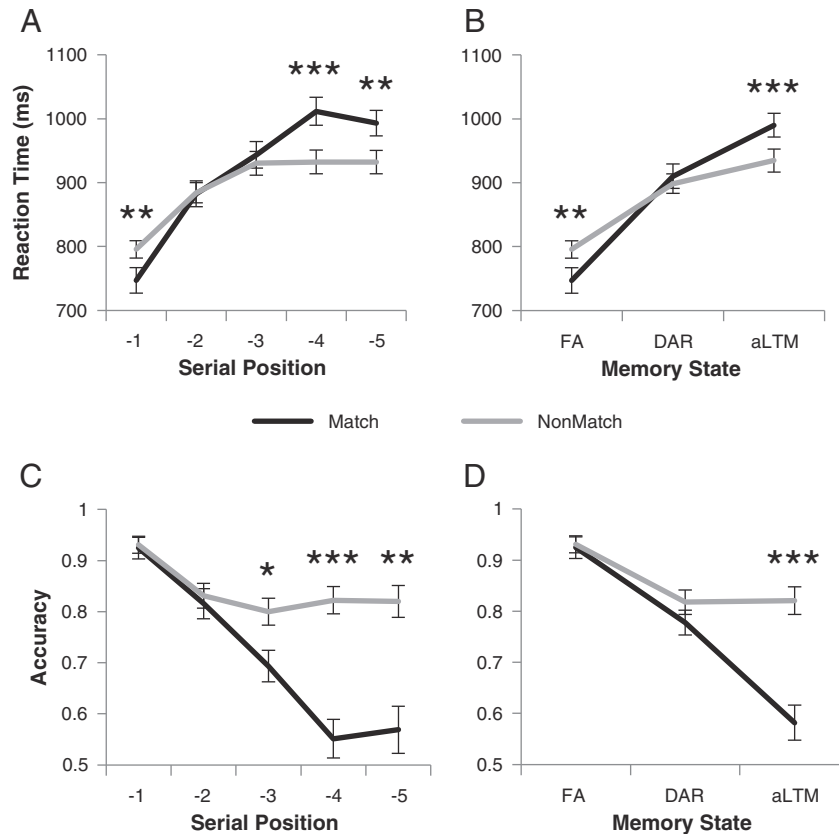
Fig. 2. Behavioral data. A) Reaction time data as a function of serial position (SP) and probe-type (match/non-match). For SP −1 probes (i.e. focus of attention), match decisions were made faster than non-match decisions. Match and non-match decisions were made equally quickly for SP −2 and −3 probes. Match decisions were made slower than non-match decisions for SPs −4 and −5. B) The same data re-grouped into hypothesized memory states as a function of individual capacity estimates. C) Accuracy data as a function of SP and probe-type. Accuracy across both probe-types dropped from SP −1 to SP −2. For non-match probes, accuracy remained constant thereafter. For match probes, accuracy continued to drop before plateauing at SP −4. D) The same data re-grouped into hypothesized states as a function of individual capacity estimates.*p<0.05; **p<0.005; ***p<0.0005. FA — focus of attention; DAR — direct access region; aLTM — activated portion of long-term memory.

Finally, we examined eye movements as a function of probe position. A one-way ANOVA of the effect of SP on eye movement found no significant effect (F(4,22) = 1.07, p > 0.35). As a result, behavioral and neural differences cannot be attributed to differences in eye movements.

Imaging results

Neural correlates of accessing the focus of attention

We began by looking for areas selectively active for accessing the FA. To that end, we searched for voxels more active for probes matching the FA relative to probes matching other states. To ensure that this contrast did not include voxels that were simply inversely related to difficulty, we constrained the analysis to voxels that did not show a difference between access to the DAR and the aLTM (see Materials and methods for full details). This analysis revealed activations in the bilateral inferior parietal cortex, situated primarily in the supramarginal gyrus, but extending dorsally to the inferior parietal lobule and posteriorly to the angular gyrus (Fig. 3A; Table 1). In accordance with previous observations, these activations did not extend into the intra-parietal sulcus, remaining distinctly ventral (Nee and Jonides, 2008, 2011). Activations were also prominent in the bilateral middle temporal gyrus with the right-sided cluster merging with the inferior parietal activations along the temporo-parietal junction. Activations were also observed in the right inferior temporal gyrus and medial frontal gyrus. A targeted analysis of the MTL did not reveal any significant activations in the MTL involved in accessing the FA. In summary, the activations in the inferior parietal and inferior temporal cortices mimicked previously documented patterns in the verbal domain. However, the activations observed here were distinctively more bilateral and more prominent along the middle temporal gyrus than observed in previous reports.

Neural correlates of accessing the direct access region

Next, we looked for areas selectively active for accessing the DAR. At our whole-brain threshold, no areas demonstrated significant selective activation for accessing the DAR. However, a targeted search of the MTL revealed activations in the right hippocampus extending into the parahippocampal gyrus/entorhinal cortex (Fig. 3B; Table 1). This result replicates previous studies implicating the MTL for retrieving information in the DAR (Nee and Jonides, 2008, 2011; Oztekin et al., 2009, 2010).

Neural correlates of accessing the activated portion of long-term memory

Finally, we looked for areas selectively active for accessing the aLTM. This analysis revealed two areas in left VLPFC (Fig. 3C; Table 1). The first was situated in the anterior insula and the second was focused in the inferior frontal gyrus, pars opercularis. This latter cluster extended anteriorly into the inferior frontal gyrus, pars triangularis, as well as posteriorly into the inferior frontal junction. Once again, this result replicated previous findings in the verbal domain (Nee and Jonides, 2008, 2011; Oztekin et al., 2009, 2010) demonstrating a role of the left VLPFC in retrieving information from the aLTM even when the information is non-verbal in nature. A targeted search within the MTL did not reveal additional activations in the MTL selective for retrieval from the aLTM.

Oztekin et al. have suggested that the MTL is involved in retrieval of information that does not reside in the FA (Oztekin et al., 2009, 2010). By this account, the MTL is involved in retrieval from both
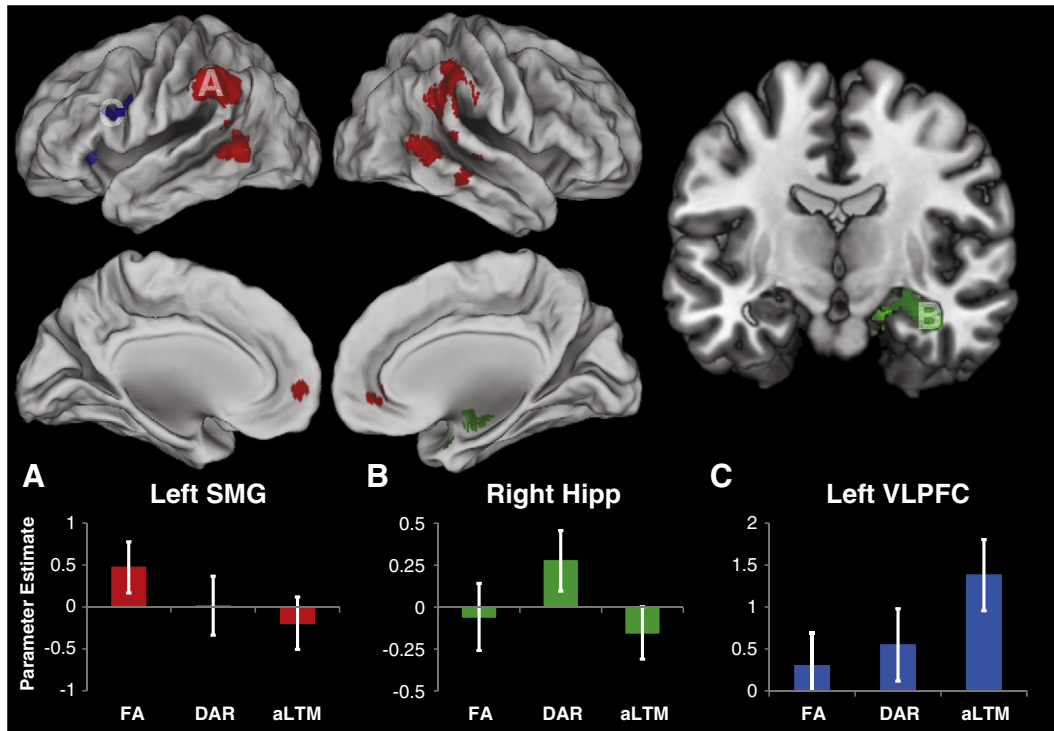
Fig. 3. Neural correlates of access to visual short-term memory. Top: areas showing significantly different activation during the access of different putative states of visual short-term memory (p<0.05 FWE corrected). Red: areas selectively active for accessing the focus of attention (FA). Green: areas selectively active for accessing the direct access region (DAR). Blue: areas selectively active for accessing the activated portion of long-term memory (aLTM). Bottom: parameter estimates drawn from statistically independent regions-of-interest (ROI) for each contrast of interest (see Supplemental Methods for ROI details). A) Parameter estimates drawn from the left supramarginal gyrus (SMG) demonstrating selectivity for accessing the FA. B) Parameter estimates drawn from the right hippocampus (hipp) demonstrating selectivity for accessing the DAR. Section cut at y = −10. C) Parameter estimates drawn from the left ventrolateral prefrontal cortex (VLPFC) demonstrating selectivity for accessing the aLTM. For more detailed visualization of these activation patterns, see Supplemental Fig. 2.

the DAR and the aLTM with activation differences for retrieval from these states reflecting differences in memory strength. In the above analysis, activations related to the access of the aLTM were contrasted against the access of both the FA and the DAR. Therefore, the inclusion of the DAR as a baseline, which elicits MTL activation, may have

Table 1
Neural correlates of access to visual short-term memory.

State	x	y	z	Extent	Z	BA	Region
FA	−66	−50	10	474	4.68	21	Left MTG
	−58	−58	2		4.43	37	Left MTG
	−66	−52	2		4.31	21	Left MTG
	−58	−40	38	605	4.49	40	Left IPL, SMG
	−58	−54	32		4.16	39	Left AG
	60	−48	4	1608	4.46	21	Right MTG
	54	−44	30		4.39	40	Right SMG, IPL
	60	−62	18		4.38	21, 22	Right MTG, STG
	64	−22	−6	163	3.85	21	Right MTG
	64	−30	−14		3.54	20	Right ITG
	54	−24	−4		3.28	21	Right MTG
	−6	54	−2	192	3.58	10	Medial FG
	0	46	−4		3.52	10	Medial FG
DAR	20	−4	−18	319	3.32		Right hipp
	36	−8	−22		2.75		Right hipp
	20	8	−26		2.59	34, 28	Right PHG
aLTM	−32	28	−2	117	4.14	13, 47	Left ant insula, IFG — orb
	−50	12	24	205	4.01	44, 45, 9	Left IFG — oper, tria

Areas selectively activated to access of each state of short-term memory. Peak activations reported in MNI space. Abbreviations: AG — angular gyrus; aLTM — activated portion of long-term memory; ant — anterior; BA — Brodmann's Area; DAR — direct access region; FA — focus of attention; FG — frontal gyrus; hipp — hippocampus; IFG — inferior frontal gyrus; IPL — inferior parietal lobule; ITG — inferior temporal gyrus; MTG — middle temporal gyrus; oper — pars opercularis; orb — pars orbitalis; PHG — parahippocampal gyrus; SMG — supramarginal gyrus; STG — superior temporal gyrus; tria — pars triangularis.

masked MTL activations related to the retrieval of information in the aLTM. To explore this possibility, we directly contrasted the access of the aLTM with the access of the FA within the MTL. Even in this targeted contrast, we did not find activations in the MTL related to the retrieval of aLTM. Hence, in the present data set, the MTL was not involved in the retrieval of information in the aLTM. We discuss the discrepancies between our data and that of Oztekin et al. in more detail in the Discussion section.

Domain general neural correlates of accessing short-term memory

The activation patterns described above closely resemble the triple dissociation we previously found that supported a 3-state model of verbal memory (Nee and Jonides, 2011). The commonality across verbal and visual domains suggests that access to qualitatively distinct memory states relies on domain general neural correlates and mechanisms. To examine this issue more directly, we performed valid conjunction analysis to reveal areas common to both verbal memory access and visual memory access. Common activation for accessing the FA was found in the left inferior parietal cortex (MNI center −56 −52 40) including the inferior parietal lobule (area 40) and angular gyrus (area 39). Common activation for accessing the DAR was found in the right hippocampus (MNI center 24 −10 −16). Finally, the left VLPFC was jointly active for accessing both verbal aLTM and visual aLTM in the left anterior insula (MNI center −32 26 2) and left inferior frontal gyrus, pars opercularis (MNI center −46 8 26). These results are summarized in Fig. 4 and Table 2.

Discussion

The present study was designed to test a 3-state model of visual STM. Using an object-based version of a serial item-recognition task that has been used to test a 3-state model of verbal STM, we found

a triple dissociation supporting a 3-state model of visual STM. Inferior parietal areas were selectively active during the access of the FA, the MTL was selectively active during the access of the DAR, and the left VLPFC was selectively active during the access of the aLTM. Furthermore, direct comparison between areas involved in accessing the 3-states of visual STM with areas involved in accessing the 3-states of verbal STM revealed overlapping regions of neural activation. These results suggest that areas involved in accessing the 3-states of memory are domain general, revealing a common STM architecture across different forms of content.

Inferior parietal cortex and the focus of attention

The inferior parietal cortex and temporo-parietal junction are central nodes in the so-called ventral attention system (Corbetta and Shulman, 2002). This system mediates reflexive attention or automatic orienting to salient events. The functions of the ventral attention system contrast with the dorsal attention system that enables goal-directed attention during search processes. While these ideas originated in the perceptual domain, it has been proposed that similar mechanisms exist in the memory domain (Cabeza et al., 2008, 2012). In memory, it has been hypothesized that the inferior parietal cortex is recruited when salient recollected episodes reflexively capture attention. Recently, these ideas have been extended to a variety of other cognitive domains in which the inferior parietal cortex is thought to mediate bottom-up attention in a domain-general manner (Cabeza et al., 2012). Here, we have demonstrated a similarly domain-general function of the inferior parietal cortex during the access of the FA in both visual STM and verbal STM.

Although we documented overlapping activations in the inferior parietal cortex for accessing the FA in both verbal STM and visual STM, the center-of-mass of these activations appears to be distinct across domains. In verbal STM, activations were situated more posteriorly in the angular gyrus (area 39) while in visual STM, activations were situated more anteriorly in the supramarginal gyrus (area 40). Similar differences have been noted when comparing inferior parietal activations across other domains (Cabeza et al., 2012). Cabeza et al. (2012) have argued that the variability in center-of-mass across domains results from differences in connectivity patterns of parietal sub-regions which reflect different sources of input to the inferior parietal cortex. Thus, while the inferior parietal cortex as a whole underlies bottom-up attention, different sub-regions of the inferior parietal cortex mediate attention to different kinds of inputs commensurate with their respective connectivity profiles.

While it is clear that the inferior parietal cortex is recruited across a variety of domains, its role is typically described as a "circuit breaker" that enables reorientation to stimuli that are not the current focus of attention. On the surface, this function appears to be at odds with the present data. Here, the inferior parietal cortex was active when a probe matched the FA. Hence, it seems that attention should have remained fixed on the same representation rather than reorienting to a new representation. This issue can be reconciled if one considers the idea of orienting attention inwards to memory versus outwards to the environment. We suggest that when a perceptual stimulus matches the FA in STM, the focus of attention is switched from memory to the environment in an automatic fashion. This idea is consistent with studies of visual search demonstrating activation of the dorsal attention system during the search process itself, but activation of the ventral attention system when a target is found (Corbetta et al., 2000). In the case of visual search, an attentional template is held in the FA in STM. When the target is found in the environment, attention is automatically drawn toward it, eliciting reorientation from the internal template to the external stimulus. Recently, it has been proposed that such attentional capture only occurs for stimuli matching the single-item FA, but not other information held in STM (Olivers et al., 2011). Hence, the FA in STM may produce a qualitatively distinct boost in the salience of external stimuli that share features with the FA, thereby facilitating automatic reorienting from memory to perception.

Medial temporal lobe and the direct access region

While the MTL is known to be involved in the formation and retrieval of LTM (Davachi, 2006; Diana et al., 2007; Eichenbaum et al., 2007; Ranganath and Ritchey, 2012), a growing literature has demonstrated the involvement of the MTL in the maintenance and retrieval of information in STM (Axmacher et al., 2008, 2010; Cabeza et al., 2002; Nee and Jonides, 2008, 2011; Nee et al., 2008; Oztekin et al., 2009, 2010; Ranganath and Blumenfeld, 2005; Ranganath and D'Esposito, 2001; Rissman et al., 2008). Although
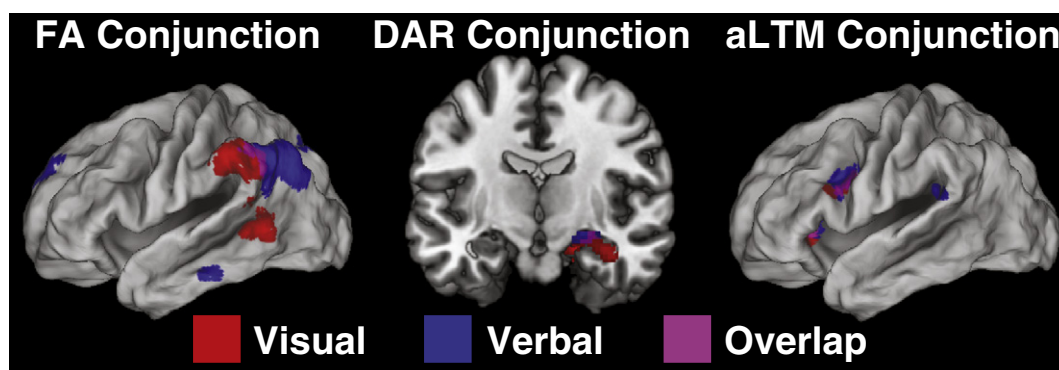
Table 2
Regions common to verbal short-term memory and visual short-term memory.

State	x	y	z	Extent	Area	Region
FA	−56	−52	40	97	40, 39	Left IPL, AG
DAR	24	−10	−16	47		Right hipp
aLTM	−32	26	2	41	13	Left ant insula
	−46	8	26	42	44	Left IFG — oper

Results of a conjunction analysis between the present study and Nee and Jonides (2011). Coordinates are reported in MNI space and reflect center-of-mass. For abbreviations, see Table 1.

Fig. 4. Conjunction results. Comparison between the present data and comparable analyses on a verbal paradigm described in Nee and Jonides (2011). Red: neural correlates of access to visual short-term memory. Blue: neural correlates of access to verbal short-term memory. Violet: overlap.

some authors have suggested that the MTL supports the retention of information over brief intervals only when STM-capacity is exceeded (Jeneson and Squire, 2012), our data demonstrate that the MTL is actually more activated when putative capacity-limits are not exceeded. Other neuroimaging studies have shown similar patterns with MTL activation being maximal within typical STM capacity-limits (Oztekin et al., 2009, 2010; Zarahn et al., 2005). However, these activation-based data are in sharp contrast with data from MTL-damaged patients who demonstrate memory impairments only for supra-capacity items (Jeneson et al., 2010, 2012). How can these data be reconciled?

The patient data indicate that the MTL may not be strictly essential for recognition-based processes of STM. Nevertheless, the MTL may play an important role in STM. We have suggested that the MTL mediates dynamic binding that inter-associates content in the DAR to the current context (Nee and Jonides, 2011). This hypothesis is consistent with contemporary theories of the hippocampus which suggest that it is involved in item–context associations, regardless of time-scale (Davachi, 2006; Diana et al., 2007; Eichenbaum et al., 2007; Ranganath and Ritchey, 2012). In short-term item-recognition tasks, the retrieval of item–context associations is non-essential. However, a probe matching an item in the DAR may nevertheless serve as a cue that triggers the retrieval of item–context associations producing the observed MTL activation. Furthermore, the degree to which the MTL is recruited during STM tasks predicts subsequent LTM formation (Ranganath et al., 2005), indicating that MTL involvement in STM does play an important function. Hence, MTL-mediated dynamic binding of items in the DAR may form the basis of new LTM encoding.

While the above proposal accounts for MTL activation related to the DAR, it does not explain why patients should be disproportionately impaired when STM-capacity is exceeded. We have suggested that supra-capacity items are held in the aLTM and are not bound to the current context. Hence, presentation of items that match information in the aLTM does not automatically trigger contextual retrieval. However, engagement of contextual retrieval processes should facilitate appropriate categorization of items in the aLTM as belonging to the currently relevant target set. To the degree that MTL-mediated contextual retrieval is engaged, healthy subjects should show superior performance to MTL-damaged patients. In the present study, subjects may have engaged contextual retrieval only minimally for probes matching items in the aLTM since the recognition probes favored familiarity-based strategies. However, other studies that have employed a two-alternative forced-choice procedure that required subjects to identify an old item against a new item have demonstrated MTL activation for retrieval from the aLTM (Oztekin et al., 2009, 2010). Those paradigms may have placed greater demands on contextual retrieval thereby eliciting MTL activation for probes matching items in the aLTM. Thus, we suggest that the MTL is necessary for contextual retrieval and that engagement of this process facilitates the appropriate contextualization of items maintained in the aLTM. However, what distinguishes the DAR and the aLTM is that contextual retrieval is automatically triggered for the former, while top-down processes may be necessary to recruit the MTL for the latter. We turn to one potential source of such top-down recruitment next.

Ventrolateral prefrontal cortex and the activated portion of long-term memory

Activation of the left VLPFC is ubiquitous in STM research. Recent meta-analyses have demonstrated that the left VLPFC is consistently active for both the storage (Rottschy et al., 2012) and executive processing of identity-based content (Nee et al., 2013). We have proposed that the left VLPFC performs a general selection function on identity-based content that enables a broad class of storage and executive processes of STM (Nee et al., 2013). In the present study,

the left VLPFC showed selective activation during the access of the aLTM. Functionally, this activation may have reflected the selection of details regarding the probe representation to determine whether the item was old or new. Previous research has demonstrated that similar left VLPFC areas are involved in evidence accumulation in the perceptual domain (Ploran et al., 2007, 2011). Furthermore, we demonstrated that during verbal STM retrieval, the left VLPFC shows functional connectivity with various sources of visual, semantic, and phonological information that may underlie evidence accumulation processes in STM (Nee and Jonides, 2008). Interestingly, we have also demonstrated that the left VLPFC shows functional connectivity with the MTL when the retrieval of contextual details is critical to resolve proactive interference (Nee et al., 2007). These data are consistent with the role of the left VLPFC in the selection of details to categorize information in the aLTM.

If the above account is correct, what is the target of evidence accumulation in visual STM? One likely candidate in the present study is the fusiform face area (FFA; Kanwisher et al., 1997). The FFA is known to preferentially represent face content and previous studies have demonstrated an important role of the FFA in the storage of faces in STM (Postle et al., 2003; Ranganath et al., 2004). While ancillary to the primary analyses of interest, we examined activations in the FFA to investigate this idea. If the FFA provides evidence regarding the old/new status of a probe, we would expect activations in the FFA to increase with increased sampling (see Supplemental Methods for details of FFA definitions). Consistent with this idea, both the right FFA and left FFA demonstrated a striking SP × probe-type interaction (Fig. 5; right FFA: $F_{(4,92)} = 5.34$, $p < 0.001$; left FFA: $F_{(4,92)} = 2.9$, $p < 0.05$). This interaction was driven by a parabolic increase in activation for match probes coupled with a flat pattern of activation for non-match probes. Notably, these results were selective to the FFA as the comparably defined parahippocampal place area (PPA) did not demonstrate a similar interaction (right PPA: $F_{(4,92)} = 1.68$, $p > 0.15$; left PPA: $F_{(4,92)} = 0.55$, $p > 0.65$). Similar effects were observed when the data were grouped into hypothesized memory states (Supplemental Fig. 3). The patterns in the FFA reflect a memory search/evidence accumulation process. Specifically, a constant, high-rate of activation would be expected for non-match probes as the FFA would be sampled until a non-match decision is reached. By contrast, match decisions regarding the FA would require virtually no sampling, decisions regarding the DAR would require an intermediate amount of sampling, and decisions regarding the aLTM would require a high degree of sampling. The dip in activation for match probes drawn from SP − 5 may reflect a primacy effect that reduces sampling demands. While the present study did not have appropriate power to explore VLPFC–FFA functional interactions, this would be an intriguing area for future research to further establish search-related interactions.

Domain generality

The close correspondence between the present study and our previous research with a verbal paradigm suggests that neural correlates and mechanisms that access the 3-states of memory are domain general. However, full investigation of this idea would require the exploration of additional domains. One important candidate to test would be spatial STM. Meta-analyses have suggested a dorsal–ventral split between location-based STM on the one hand and identity-based STM on the other (Nee et al., 2013; Rottschy et al., 2012; Wager and Smith, 2003). We have proposed that selection mechanisms involved in STM form a "what"/"where" distinction which would predict that, in the least, selection mechanisms that elicit left VLPFC activations in the present paradigm would instead recruit dorsal frontal areas along the superior frontal sulcus for spatial content (Nee et al., 2013). This account predicts a "what"/"where" distinction for access to the aLTM. However, whether
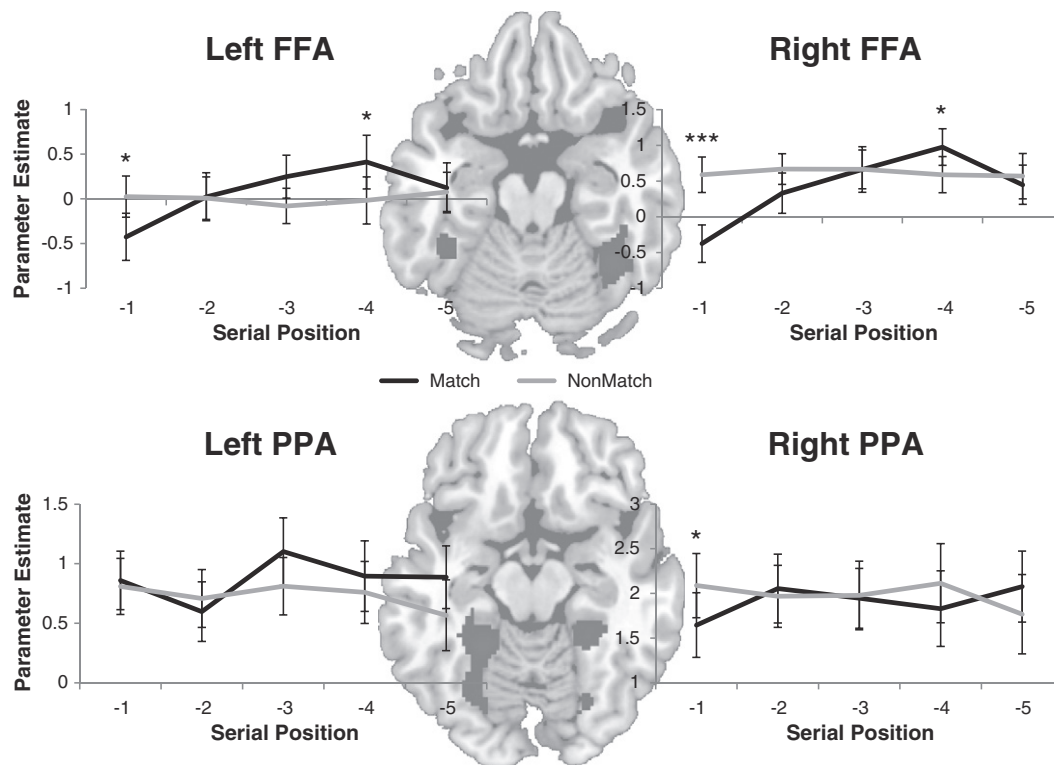
Fig. 5. Results in object-selective areas. The underlays depict face- and place-selective voxels (black) from a representative subject. Top: group-averaged activations drawn from the left and right fusiform face areas (FFAs) as a function of serial position and probe-type. Bottom: group-averaged activations drawn from the right and left parahippocampal place areas (PPAs) as a function of serial position and probe-type. *p<0.05; **p<0.005; ***p<0.0005.

access to the FA and the DAR show "what"/"where" distinctions is unclear. Such investigation would be an important next step to identify putative domain general mechanisms of STM.

Representation in visual STM

The present study used neural correlates of access to different putative states to investigate a 3-state model of visual STM. While this logic has proven useful, recent advances in neuroimaging have enabled new ways of investigating the representations of STM more directly. In a recent elegant study, Lewis-Peacock et al. utilized multi-variate pattern analysis (Norman et al., 2006) to contrast representations of the FA with non-focused content (Lewis-Peacock et al., 2012). In their task, subjects were presented with two different categories of information on each trial (e.g. visuo-spatial, phonological, semantic). After a retention interval, subjects were presented with a cue that indicated that one of the categories would be relevant for an upcoming decision. It was assumed that the cued item would be held in the FA while the uncued item would be maintained in a different state. By using a machine learning classifier, the authors found that the cued representation could be accurately decoded from the pattern of activation across the brain, while the uncued representation could not. However, this pattern reversed if a second cue switched the FA to the originally uncued item. While it is unclear whether uncued items were retained in the DAR or the aLTM, this study suggests that only the FA achieves full neural representation throughout the brain. Based on these findings, we speculate that items in the DAR may not be fully represented by active neural firing across the brain, but instead may be activated only through the MTL-mediated context. Examination of this idea would be an important avenue for future research.

Another way in which visual STM is traditionally explored is through the examination of capacity-limits. In these types of paradigms, researchers have examined neural regions that mirror individual differences in visual STM capacity. Data from both ERP (Vogel and Machizawa, 2004; Vogel et al., 2005) and fMRI (Todd and Marois, 2004; Xu and Chun, 2006) have demonstrated that the intra-parietal sulcus[2] closely tracks visual STM capacity-limits. Recent data have indicated that capacity-related intra-parietal sulcus activations are domain general across different forms of content and input modalities (Cowan et al., 2011). Such limits likely reflect the capacity of the DAR and indicate that functions of the intra-parietal sulcus may constrain how many items can be held in the DAR. An interesting future pursuit would be to examine relationships between the intra-parietal sulcus and the MTL during the storage and retrieval of representations in the DAR.

Limitations

While we have argued that the data provide evidence for a 3-state model of visual STM, it is important to note that these conclusions rest upon the assumptions of the present paradigm. In particular, we have assumed that the most recently presented item resides in the FA, items up the capacity limit reside in the DAR, and supra-capacity items reside in the aLTM. These assumptions were made in keeping with previous research and fit well with the behavioral data. However,

---

[2] Notably, these dorsal parietal regions are distinct from the ventral parietal regions related to the FA (Nee and Jonides, 2008), and the two should not be confused (see the discussion of the ventral versus dorsal parietal systems above: Inferior Parietal Cortex and the Focus of Attention).

convergent support from other paradigms is needed to solidify the 3-state model of visual STM. For example, additional support may be garnered from paradigms that require subjects to flexibly shift information among the three putative states (Lewis-Peacock et al., 2012; Oberauer, 2002). Such investigations would be an important next step in research on the architecture of STM.

Conclusion

A 3-state model of STM has proven to be an accurate description of verbal STM, unifying distinct and conflicting 2-state theories into a single account. We have demonstrated that this model also provides an accurate description of visual STM, revealing qualitative distinctions along traditional lines of attention, STM, and LTM. Despite these qualitative state distinctions, we found areas involved in the access of each state that generalize across different forms of content and that connect with various literatures outside the domain of STM. The inferior parietal cortex mediates automatic access to the FA consistent with its role in bottom-up attention. The MTL mediates access to the DAR consistent with its role in dynamically binding items to contexts. Finally, the left VLPFC mediates access to the aLTM consistent with its role in top-down selection of identity-based content. Collectively, these data indicate that models of cognition would benefit from the partitioning of memory into three distinct states each with its own distinct form of access.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.02.019.

References

Ashburner, J., Friston, K., 1997. Multimodal image coregistration and partitioning—a unified framework. Neuroimage 6, 209–217.

Axmacher, N., Schmitz, D.P., Wagner, T., Elger, C.E., Fell, J., 2008. Interactions between medial temporal lobe, prefrontal cortex, and inferior temporal regions during visual working memory: a combined intracranial EEG and functional magnetic resonance imaging study. J. Neurosci. 28, 7304–7312.

Axmacher, N., Henseler, M.M., Jensen, O., Weinreich, I., Elger, C.E., Fell, J., 2010. Cross-frequency coupling supports multi-item working memory in the human hippocampus. Proc. Natl. Acad. Sci. U. S. A. 107, 3228–3233.

Baddeley, A.D., 1986. Working Memory. Oxford University Press, Oxford.

Baddeley, A., 2003. Working memory: looking back and looking forward. Nat. Rev. Neurosci. 4, 829–839.

Baddeley, A.D., Hitch, G.J., 1974. Working memory. In: Bower, G.H. (Ed.), The Psychology of Learning and Motivation. Academic Press, New York, pp. 47–89.

Broadbent, D.E., Broadbent, M.H.P., 1981. Recency effects in visual memory. Q. J. Exp. Psychol. 33, 1–15.

Cabeza, R., Dolcos, F., Graham, R., Nyberg, L., 2002. Similarities and differences in the neural correlates of episodic memory retrieval and working memory. Neuroimage 16, 317–330.

Cabeza, R., Ciaramelli, E., Olson, I.R., Moscovitch, M., 2008. The parietal cortex and episodic memory: an attentional account. Nat. Rev. Neurosci. 9, 613–625.

Cabeza, R., Ciaramelli, E., Moscovitch, M., 2012. Cognitive contributions of the ventral parietal cortex: an integrative theoretical account. Trends Cogn. Sci. 16, 338–352.

Carpenter, P.A., Just, M.A., Shell, P., 1990. What one intelligence-test measures—a theoretical account of the processing in the Raven Progressive Matrices test. Psychol. Rev. 97, 404–431.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3, 201–215.

Corbetta, M., Kincade, J.M., Ollinger, J.M., McAvoy, M.P., Shulman, G.L., 2000. Voluntary orienting is dissociated from target detection in human posterior parietal cortex. Nat. Neurosci. 3, 292–297.

Cowan, N., 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav. Brain Sci. 24, 87–114 (discussion 114–185).

Cowan, N., Li, D., Moffitt, A., Becker, T.M., Martin, E.A., Saults, J.S., Christ, S.E., 2011. A neural region of abstract working memory. J. Cogn. Neurosci. 23, 2852–2863.

Daneman, M., Carpenter, P.A., 1980. Individual differences in working memory and reading. J. Verbal Learn. Verbal Behav. 19, 450–466.

Daneman, M., Merikle, P.M., 1996. Working memory and language comprehension: a meta-analysis. Psychon. Bull. Rev. 3, 422–433.

Davachi, L., 2006. Item, context and relational episodic encoding in humans. Curr. Opin. Neurobiol. 16, 693–700.

Diana, R.A., Yonelinas, A.P., Ranganath, C., 2007. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends Cogn. Sci. 11, 379–386.

Eichenbaum, H., Yonelinas, A., Ranganath, C., 2007. The medial temporal lobe and recognition memory. Annu. Rev. Neurosci. 30, 123–152.

Hanna, A., Loftus, G., 1993. A model for conceptual processing of naturalistic scenes. Can. J. Exp. Psychol. 47, 548–569.

Hay, D.C., Smyth, M.M., Hitch, G.J., Horton, N.J., 2007. Serial position effects in short-term visual memory: a SIMPLE explanation? Mem. Cognit. 35, 176–190.

Jaeggi, S.M., Buschkuehl, M., Jonides, J., Perrig, W.J., 2008. Improving fluid intelligence with training on working memory. Proc. Natl. Acad. Sci. U. S. A. 105, 6829–6833.

Jeneson, A., Squire, L.R., 2012. Working memory, long-term memory, and medial temporal lobe function. Learn. Mem. 19, 15–25.

Jeneson, A., Mauldin, K.N., Squire, L.R., 2010. Intact working memory for relational information after medial temporal lobe damage. J. Neurosci. 30, 13624–13629.

Jeneson, A., Wixted, J.T., Hopkins, R.O., Squire, L.R., 2012. Visual working memory capacity and the medial temporal lobe. J. Neurosci. 32, 3584–3589.

Jonides, J., Lacey, S., Nee, D., 2005. Processes of working memory in mind and brain. Curr. Dir. Psychol. Sci. 14, 2–5.

Jonides, J., Lewis, R.L., Nee, D.E., Lustig, C.A., Berman, M.G., Moore, K.S., 2008. The mind and brain of short-term memory. Annu. Rev. Psychol. 59, 193–224.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Kerr, J.R., Avons, S.E., Ward, G., 1999. The effect of retention interval on serial position curves for item recognition of visual patterns and faces. J. Exp. Psychol. Learn. Mem. Cogn. 25, 1475–1494.

Lewis-Peacock, J.A., Drysdale, A.T., Oberauer, K., Postle, B.R., 2012. Neural evidence for a distinction between short-term memory and the focus of attention. J. Cogn. Neurosci. 24, 61–79.

Luck, S.J., Vogel, E.K., 1997. The capacity of visual working memory for features and conjunctions. Nature 390, 279–281.

Lund, T.E., Norgaard, M.D., Rostrup, E., Rowe, J.B., Paulson, O.B., 2005. Motion or activity: their role in intra- and inter-subject variation in fMRI. Neuroimage 26, 960–964.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage 19, 1233–1239.

McElree, B., 2006. Accessing recent events. In: Ross, B. (Ed.), The Psychology of Learning and Motivation. Academic Press, San Diego, CA, pp. 155–200.

McElree, B., Dosher, B.A., 1989. Serial position and set size in short-term memory: the time course of recognition. J. Exp. Psychol. Gen. 118, 346–373.

Miller, G.A., 1956. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol. Rev. 63, 81–97.

Nee, D.E., Jonides, J., 2008. Neural correlates of access to short-term memory. Proc. Natl. Acad. Sci. U. S. A. 105, 14228–14233.

Nee, D.E., Jonides, J., 2011. Dissociable contributions of prefrontal cortex and the hippocampus to short-term memory: evidence for a 3-state model of memory. Neuroimage 54, 1540–1548.

Nee, D.E., Jonides, J., Berman, M.G., 2007. Neural mechanisms of proactive interference-resolution. Neuroimage 38, 740–751.

Nee, D.E., Berman, M.G., Moore, K.S., Jonides, J., 2008. Neuroscientific evidence about the distinction between short- and long-term memory. Curr. Dir. Psychol. Sci. 17, 102–106.

Nee, D.E., Brown, J.W., Askren, M.K., Berman, M.G., Demiralp, E., Krawitz, A., Jonides, J., 2013. A meta-analysis of executive components of working memory. Cereb. Cortex 23, 264–282.

Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B., 2005. Valid conjunction inference with the minimum statistic. Neuroimage 25, 653–660.

Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. 10, 424–430.

Oberauer, K., 2002. Access to information in working memory: exploring the focus of attention. J. Exp. Psychol. Learn. Mem. Cogn. 28, 411–421.

Olivers, C.N., Peters, J., Houtkamp, R., Roelfsema, P.R., 2011. Different states in visual working memory: when it guides attention and when it does not. Trends Cogn. Sci. 15, 327–334.

Oztekin, I., McElree, B., Staresina, B.P., Davachi, L., 2009. Working memory retrieval: contributions of the left prefrontal cortex, the left posterior parietal cortex, and the hippocampus. J. Cogn. Neurosci. 21, 581–593.

Oztekin, I., Davachi, L., McElree, B., 2010. Are representations in working memory distinct from representations in long-term memory? Neural evidence in support of a single store. Psychol. Sci. 21, 1123–1133.

Phillips, W.A., Christie, D.F.M., 1977. Components of visual memory. Q. J. Exp. Psychol. 29, 117–133.

Ploran, E.J., Nelson, S.M., Velanova, K., Donaldson, D.I., Petersen, S.E., Wheeler, M.E., 2007. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. J. Neurosci. 27, 11912–11924.

Ploran, E.J., Tremel, J.J., Nelson, S.M., Wheeler, M.E., 2011. High quality but limited quantity perceptual evidence produces neural accumulation in frontal and parietal cortex. Cereb. Cortex 21, 2650–2662.

Postle, B.R., 2006. Working memory as an emergent property of the mind and brain. Neuroscience 139, 23–38.

Postle, B.R., Druzgal, T.J., D'Esposito, M., 2003. Seeking the neural substrates of visual working memory storage. Cortex 39, 927–946.

Ranganath, C., Blumenfeld, R.S., 2005. Doubts about double dissociations between short- and long-term memory. Trends Cogn. Sci. 9, 374–380.

Ranganath, C., D'Esposito, M., 2001. Medial temporal lobe activity associated with active maintenance of novel information. Neuron 31, 865–873.

Ranganath, C., Ritchey, M., 2012. Two cortical systems for memory-guided behaviour. Nat. Rev. Neurosci. 13, 713–726.

Ranganath, C., DeGutis, J., D'Esposito, M., 2004. Category-specific modulation of inferior temporal activity during working memory encoding and maintenance. Brain research. Cogn. Brain Res. 20, 37–45.

Ranganath, C., Cohen, M.X., Brozinsky, C.J., 2005. Working memory maintenance contributes to long-term memory formation: neural and behavioral evidence. J. Cogn. Neurosci. 17, 994–1010.

Rissman, J., Gazzaley, A., D'Esposito, M., 2008. Dynamic adjustments in prefrontal, hippocampal, and inferior temporal interactions with increasing visual working memory load. Cereb. Cortex 18, 1618–1629.

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A.R., Schulz, J.B., Fox, P.T., Eickhoff, S.B., 2012. Modelling neural correlates of working memory: a coordinate-based meta-analysis. Neuroimage 60, 830–846.

Todd, J.J., Marois, R., 2004. Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428, 751–754.

Vogel, E.K., Machizawa, M.G., 2004. Neural activity predicts individual differences in visual working memory capacity. Nature 428, 748–751.

Vogel, E.K., McCollough, A.W., Machizawa, M.G., 2005. Neural measures reveal individual differences in controlling access to working memory. Nature 438, 500–503.

Wager, T.D., Smith, E.E., 2003. Neuroimaging studies of working memory: a meta-analysis. Cogn. Affect. Behav. Neurosci. 3, 255–274.

Xu, Y., Chun, M.M., 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. Nature 440, 91–95.

Zarahn, E., Rakitin, B., Abela, D., Flynn, J., Stern, Y., 2005. Positive evidence against human hippocampal involvement in working memory maintenance of familiar stimuli. Cereb. Cortex 15, 303–316.