

**Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks**

Daniel J. Mitchell<sup>1</sup> and Rhodri Cusack<sup>1</sup>

<sup>1</sup>MRC Cognition and Brain Sciences Unit (CBU), Cambridge, UK

Corresponding author:

Daniel J. Mitchell

MRC Cognition and Brain Sciences Unit

15 Chaucer Road

Cambridge

CB2 7EF

England, UK

[daniel.mitchell@mrc-cbu.cam.ac.uk](mailto:daniel.mitchell@mrc-cbu.cam.ac.uk)

Telephone: +44 1223 355294

Fax: +44 1223 359062

Running title:

Capacity-limited perceptual representations in human cortex

## **Abstract**

It has recently been shown, using functional magnetic resonance imaging (fMRI) with a change detection paradigm, that activity in posterior parietal cortex (PPC) correlates with the limited number of objects held in visual short-term memory (VSTM). We replicate this finding and extend it to tasks that use similar stimuli, but without explicit memory requirements. As well as a perceptual task used previously (detecting an item at fixation; Todd and Marois 2004), two additional tasks were designed to increase attentional demands across space (searching for a red item anywhere in the array) and across both space and time (detecting a staggered offset after prolonged viewing of the array). During the VSTM task, a capacity-limited set-size effect was seen in PPC as well as occipital and frontal regions. However, the PPC showed similar activity during two of the tasks not requiring VSTM. These findings cannot easily be explained by behavioural performance measures or memory demands alone, suggesting a role of the PPC in processing a limited number of discrete object representations, whether in the current perceptual scene or working memory. The differential influence of item load across perceptual tasks is consistent with task requirements affecting the form of these representations.

**Keywords:** fMRI, human, parietal lobe, perceptual organisation, visual attention, working memory

## Introduction

VSTM is the ability to maintain a small amount of visual information for a period of seconds, in contrast to sensory or ‘iconic’ memory (IM) which stores a richer, but more fragile, representation that decays within half a second and is more sensitive to eye movements and intervening distracters (Phillips 1974). VSTM has a capacity limit of approximately four items (Cowan 2001), although this varies somewhat between individuals (Todd and Marois 2005; Vogel and Machizawa 2004), and depends on the particular task and objects being remembered (Alvarez and Cavanagh 2004; Xu and Chun 2006). The number of items held in working memory ( $K$ ) can be estimated in a change detection task using a formula introduced by Pashler (1988) and modified by Cowan (2001)(see methods). Although this measure is typically interpreted as  $K$  items being perfectly encoded, with no encoding of any other items, it is possible that a larger number of lower fidelity representations are being stored (Wilken and Ma 2004). It might therefore be useful to consider  $K$  as the number of “object’s-worth” of information that can be maintained (Vogel and others 2006).

There is still debate as to the nature of the representations that are stored in working memory. It has been demonstrated that whole, multi-feature objects can be successfully encoded almost as well as single feature objects, but that performance declines if the same number of features are distributed amongst multiple objects (Luck and Vogel 1997). This suggests that VSTM is limited by the number of perceptual objects rather than by simple features, and is consistent with object effects seen in numerous other situations of which a small selection include selective attention to overlapping stimuli (Duncan 1984; O’Craven and others 1999), multiple object tracking (MOT) (Scholl and others 2001; Yantis 1992), inhibition of return (Paul and Tipper 2003), saccade planning (Vergilino-Perez and Findlay 2006), and target detection in patients with unilateral neglect (Humphreys 1998; Tipper and Behrmann 1996). Such object effects may be mediated by constructs such as object files which store or index information about discrete spatiotemporally continuous entities (Kahneman and others 1992). However, it has been variously suggested that VSTM also stores relational information (Jiang and others 2000; Jiang and others 2004), is limited by both the number of objects and their

feature complexity (Alvarez and Cavanagh 2004; Awh and others 2007; Olson and Jiang 2002; Xu and Chun 2006), or is limited by object parts and feature conjunctions rather than the number of objects *per se* (Davis and Holmes 2005). Since most working memory tasks involve multiple component process, both object-based and feature-based processing may occur, but for different components of the task. For example, an EEG study which varied both object load and feature load suggested that encoding and retention are object-based, while retrieval of the object to compare with a probe must be done on a feature-by-feature basis (Busch and Herrmann 2003).

Although the frontal lobe has long been associated with working memory processes, a number of studies suggest that frontal regions contribute more to strategic or goal-based components, while the parietal lobe may be more important for maintaining the limited contents of VSTM (Bor and others 2003; Ferreira and others 1998; Linden and others 2003; Owen 2004; Pochon and others 2001; Todd and Marois 2004; Vogel and Machizawa 2004; Xu and Chun 2006). It was recently demonstrated using functional magnetic resonance imaging (fMRI) that the activity of a bilateral region in human posterior parietal cortex (PPC) correlates with the number of objects encoded in VSTM (Todd and Marois 2004). The authors took advantage of the capacity limit by estimating  $K$  over a range of set-sizes (from one to eight items) during a delayed visual matching-to-sample task, and identifying brain regions whose activity correlated with this behavioural measure. A bilateral PPC region (near the intersection of the intraparietal sulci (IPS) and intraoccipital sulci (IOS)) showed such an effect, with the blood-oxygen-level-dependent (BOLD) signal increasing from set-sizes one to four, but levelling off thereafter. In a series of experiments, it was shown that this effect generalised to other types of visual stimuli, was caused by neither haemodynamic saturation nor insufficient time to encode the items, and was unlikely to be explained by task difficulty (task difficulty increased linearly with set-size, but the BOLD activity was explained significantly better by a quadratic function than by a linear function). It thus seems that PPC is involved in the storage and/or processing of the limited number of mental objects that can be represented in VSTM. This suggestion is bolstered by observations that differences in the BOLD activity of this region across

different subjects also correlate with individual differences in their short-term memory capacity (Todd and Marois 2005). A subsequent study (Xu and Chun 2006) extended this finding to show that a superior part of the IPS and the lateral occipital complex (LOC) track the number of objects stored, largely independent of spatial location, but also limited by the complexity of the objects. In contrast, a more inferior region of the IPS reached asymptote around set-size four regardless of object complexity, but only when the objects occupied different spatial locations.

Electrophysiological measures provide converging evidence of limited capacity neural activity during VSTM tasks. For example, interhemispheric difference waveforms, such as the ‘contralateral delay activity’ to an attended memory array, asymptote as VSTM capacity is reached, and also correlate with individual differences in VSTM capacity. This effect was found to predominate over posterior parietal and lateral occipital electrodes, consistent with the fMRI data (Vogel and Machizawa 2004).

Todd and Marois also conducted a control ‘iconic memory’ (IM) experiment which was identical to their VSTM experiment, except that only the sample display was presented and rather than performing a memory task, the subjects simply had to report the presence or absence of an item in the centre of the display. It was found that the BOLD signal in IPS/IOS was attenuated and was no longer significantly affected by set-size. It was therefore concluded that “the IPS/IOS is insensitive to the perceptual load of the visual scene” (Todd and Marois 2004). However, various other studies have explicitly or implicitly suggested that activity in similar posterior parietal regions is modulated by the number of perceived objects in the absence of any obvious working memory requirements beyond the time span of sensory memory. Examples can be found from across a range of different paradigms and even sensory modalities, including judgements of numerosity (Piazza and others 2004), the perception of ambiguous auditory stimuli (Cusack 2005), and the perception of visual objects defined by illusory contours (Ritzl and others 2003). It has been argued that “the short-term retention of information is supported by sustained activity in the same brain regions that are

responsible for the processing and representation of this information in non-working memory situations, such as perception” (Postle 2006). If this view is correct, then it might be expected that areas sensitive to the number of items held in working memory would also be sensitive to the number of items processed in non-working memory situations and vice-versa. Furthermore, the capacity limitations which are so evident in working memory paradigms may reflect constraints on the processing of underlying perceptual representations. Whether or not the IPS shows sensitivity to perceptual load may depend on whether the items concerned are being attended. For example, the appearance of new objects leads to an increase in IPS activity, even if they are task-irrelevant, but only when they are being attended (Hon and others 2006). Several studies have shown increased IPS activity as the number of attended objects increases, using a multiple-object tracking task (Culham and others 2001; Jovicich and others 2001; Tomasi and others 2004). This suggestion would be consistent with proposals that VSTM capacity is limited by how much information can reside at one time in the focus of attention (Cowan 2001) and that the role of the IPS in VSTM is likely to be attention-based (Marois and Ivanoff 2005). On the other hand, parietal sensitivity to presented objects does not seem to require conscious awareness of their presence (Fang and He 2005).

Although many of the imaging studies mentioned above have not focused on capacity limits, it has been frequently demonstrated that our capacity for perception and parallel deployment of attention can be severely limited in tasks where working memory, at least as commonly conceived, does not play an obvious role. For example, the number of orientation-defined transparent-motion dot surfaces that can be simultaneously perceived is approximately two. This limit can be raised by increasing the salience of the motion signals, however there still appears to be a higher-order upper limit of three surfaces which is insensitive to manipulations of signal strength (Greenwood and Edwards 2006). The number of independently moving objects that can be simultaneously tracked is typically reported as being limited to approximately three to five in a single plane (Culham and others 2001; Jovicich and others 2001; Pylyshyn Z. W. and Storm 1988; Tomasi and others 2004), although this may be increased to as many as seven when the targets and distracters are distributed

across multiple depth planes (Viswanathan and Mingolla 2002). When people are asked to report the numerosity of a set of briefly presented objects, they are accurate and confident for sets of below about three to six items and reaction times increase gradually with set-size. As set-sizes increase beyond this range, all three functions show apparent discontinuities, with accuracy and confidence dropping and an accompanying sudden increase in reaction times. The different methods of judgement either side of this discontinuity are termed ‘subitizing’ and ‘estimating’ respectively (Kaufman and Lord 1949; Trick and Pylyshyn 1994). Even at only 22 weeks of age, a similar limit on the sensitivity to object number is apparent, with infants being able to discriminate two from three, but not four from six (Starkey and Cooper 1980). Illusory motion of a line away from a location which has been cued immediately before (Hikosaka and others 1993) has also been used to show that multiple items can be pre-cued in parallel, but only up to a limit of approximately three to seven (Schmidt and others 1998). The origin of the capacity limits in these tasks is often viewed as attentional or pre-attentive rather than memory-based and may share a common mechanism (Pylyshyn Z. W. 2001; Trick and Pylyshyn 1994). In summary, therefore, there is evidence for a parallel, capacity-limited process in early visual perception and the capacity limit is found to be in a similar range across various paradigms and similar to the capacity limit observed in working memory tasks.

The purpose of the current study was to explore how the modulation of BOLD activity by visual object load depends on the specific task requirements, and whether such sensitivity in perceptual tasks reflects the same capacity limitations apparent in working memory situations. To this end, we aimed to replicate the findings of Todd and Marois (2004) described above, and to add two further perceptual controls, one which forced the subjects to attend to all of the items presented, rather than just the central position (responding ‘yes’ or ‘no’ depending on whether or not a red disc was present), and another which encouraged the subjects to attend to the items for a period of time comparable to the memory delay in the original condition (responding ‘yes’ or ‘no’ at the offset of the display, depending on whether one of the discs faded away over 100ms rather than vanishing

instantaneously). It was hypothesised that if the regions whose activity correlates with the limited number of items stored in VSTM are insensitive to the perceptual load of the visual scene *per se*, then their activity should be unrelated to the number of presented items in any of the perceptual tasks. On the other hand, if they are sensitive to the number of objects being attended, then set-size related activity might be revealed in the additional control tasks by increasing the attentional demands across space and time. If any such activity has no capacity limit (or a limit higher than the maximum set-size of eight items used here) then it might be expected to increase linearly with set-size, whereas if it has a capacity limit reflecting that of VSTM then it should be preferentially revealed by an asymptotic regressor derived from the behavioural *K* function.

## **Materials and Methods**

### *Participants.*

16 healthy subjects (aged 18-36 years, mean 22.8; 10 female; DJM and 15 naïve volunteers) took part. All reported normal colour vision and visual acuity, gave their informed, written consent and were reimbursed for their time. The behavioural data of all subjects were analysed, although four subjects were discarded prior to statistical analysis of the fMRI data, two for excessive movement during the fMRI acquisitions (frequently more than 3mm), and two because no significant activity could be detected in a control contrast of right hand vs. left hand motor activity. The behavioural data presented is from the remaining twelve subjects only.

### *Procedure.*

The task and stimulus presentation (fig. 1) for the VSTM and IM conditions were as close as possible to those used by Todd and Marois (2004), although to reduce scanning time only three set-sizes (1, 4 and 8) were used and the null events were replaced by variable duration inter-trial-intervals (ITI). In all conditions, two spoken digits were initially presented serially over headphones, for 250ms each, followed by a silent period and auditory mask of similar durations. These were to be retained for the duration of the trial, in order to occupy auditory/verbal working memory systems



and so prevent verbal recoding/rehearsal of the visually presented discs (Baddeley 1986; Todd and Marois 2004). A fixation point was then presented alone for 1500ms and then replaced by an array of 1, 4 or 8 different coloured discs (each  $0.38^\circ$  in diameter). The discs were selected from a palette of ten colours (red, orange, yellow, light green, dark green, light blue, dark blue, pink, white and black), randomly positioned in a 3x3 matrix ( $1.38^\circ \times 1.38^\circ$ ) and centred at fixation. In the VSTM task, this sample display lasted for 150ms, then, after a retention interval of 1200ms, a single probe disc reappeared and subjects were asked to judge whether or not a disc of this colour had occupied the same position in the original display. Probes were always presented at an occupied position in the sample display and matched the sample colour on a random 50% of trials. For non-match trials at the larger set-sizes, the probe colour had a 50% chance of having occurred elsewhere in the sample display. The sequence was repeated 96 times, 32 for each set-size, with the set-sizes being randomly intermixed. Working memory load was estimated for each set-size using Cowan's formula:  $K=N(H-F)$ , where N is the set-size, H is the hit rate and F is the false alarm rate (Cowan 2001). Sensitivity ( $d'$ ) was estimated using signal detection theory, as a measure of performance independent of any response bias that might vary across the different set-sizes.

The IM condition was identical to the VSTM condition except that the probe disc did not appear and the subjects responded immediately to the sample display according to whether or not a disc was present at the fixation position.

The 'Extended Spatial Attention' (ESA) condition was identical to the IM condition except that subjects had to indicate whether or not a red disc was present anywhere in the display.

The 'Extended Spatial and Temporal Attention' (ESTA) condition was identical to the IM and ESA conditions except that the sample display was presented for between 750ms and 1350ms and the subjects had to decide whether or not one of the discs faded away over 100ms prior to the offset of

the whole array. The exact offset time was jittered within the 600ms range to encourage subjects to maintain attention throughout.

Finally, at the end of each trial, in all conditions, a pair of digits was presented visually and the subjects had to indicate whether or not this was the same pair as presented at the beginning of the trial. The four conditions were run as separate blocks, the order counterbalanced across subjects, with the constraint that the VSTM and IM tasks were adjacent, as were the ESA and ESTA tasks. Subjects received practice at all four tasks prior to entering the scanner, until they and the experimenter were satisfied that they understood the instructions. Subjects were asked to respond as quickly and as accurately as possible in all conditions, although accuracy was emphasised. Yes/no responses were made with the index fingers of each hand using a button box. Response mapping was counterbalanced across subjects.

#### *fMRI data acquisition and analysis*

Volunteers were scanned at the MRC Cognition and Brain Sciences Unit, Cambridge, UK, on a Siemens TIM Trio 3T machine. The tasks were performed during four separate sessions of echoplanar imaging with 439 volumes acquired in each (including 12 seconds of initial dummy scans which were discarded to allow for equilibration of the signal). Acquisition parameters were as follows: TR = TA = 2s; TE = 30ms; flip angle = 78°; thirty two, 3mm thick, interleaved slices of matrix size 64x64 with a 25% gap; approximate voxel size 3x3x3.75mm. Visual Basic 6, running under Windows XP, was used to create and present the stimuli as well as to collect the responses. Visual stimuli were projected onto a screen using a Christie video projector with a 60Hz refresh rate, and viewed using an angled mirror positioned approximately 90mm above the subjects' eyes. Auditory stimuli were presented binaurally over Siemens tube headphones at a volume comfortable for the volunteer.

Functional data were pre-processed and analysed in a standard manner using SPM 5 (Wellcome Department of Imaging Neuroscience, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>). Pre-processing included slice-time correction, realignment to a reference image, non-linear normalisation to the MNI (Montreal Neurological Institute) template brain, and spatial filtering with a 10mm full-width half-maximum Gaussian kernel. This was automated using in-house software (<http://imaging.mrc-cbu.cam.ac.uk/imaging/AutomaticAnalysisManual>).

The measured signal time-series from each voxel was fitted to a combination of regressors under a general linear model. The first two regressors modelled the neural response to left and right hand motor responses, while the third regressor modelled any set-size independent effect of each visual sample-display event. In each case, the events of interest were convolved with a canonical haemodynamic response function to generate a timecourse of predicted neural activity associated with each event type. The next two regressors, of primary interest, were similar to the sample-display regressor except that each event was weighted, prior to convolution, either by the number of items in the display, or by the estimated number of items stored in memory ( $K$ ) at that set-size. This analysis allowed the BOLD signal to be partitioned at the first-level model into components that varied linearly with set-size or plateaued with the number of items in VSTM. A second model was run to determine the response to individual set-sizes, in which the three regressors modelling different weightings of all sample displays were replaced by three new regressors each modelling displays of a single set-size. In a supplementary analysis, an extra regressor was included to model those trials where the probe disc was taken from the sample colour set, but occurred at a different position. All analyses also included regressors to model any mean effect of each session, the realignment parameters as unconvolved regressors to model out residual effects of subject movement, and binary regressors to model out occasional scans where a striping artefact was detected. (These were caused by a combination of subject movement with a sequence that used thin, interleaved slices.) The data and the model were high-pass filtered with a cut-off at 200s to remove low frequency drift in the signal. Contrasts were run to estimate the signal change associated with

these regressors at each voxel, and the resulting maps from each subject were entered into a second level analysis treating subjects as a random effect. Statistical parametric maps of t-statistics were thresholded at a significance level of  $p < 0.05$ , corrected for multiple comparisons using the false discovery rate (FDR). All experimental sessions (i.e the VSTM, IM, ESA and ESTA tasks) were treated similarly.

In addition to this whole-brain approach, regions of interest (ROIs) were defined as spheres of 10mm radius, centred on the Talairach coordinates reported by Todd and Marois (2004): left IPS: -22, -65, 42; right IPS: 23, -59, 45; left VO: -34, -68, 0; right VO: 32, -77, -9; ACC: 8, 19, 30. MarsBar (<http://marsbar.sourceforge.net/>) was used to extract mean BOLD signal time-courses from these regions, and to estimate the coefficients associated with the weighted regressors in the first analysis, and the signal change associated with the sample display at each set-size in the second analysis.

## Results

### *Behavioural*

In all comparisons of performance, the  $d'$  and percentage correct measures led to identical conclusions.  $d'$  is reported except where percentage correct is useful for comparison with previous results.

Behavioural performance on the VSTM task (fig. 2) was comparable to that reported by Todd and Marois (2004):

- Performance on the flanking verbal memory test was high (mean 99% correct across all set-sizes) and independent of set-size (Percent correct:  $F(2,22)=0.647$ ,  $p > 0.05$ );  $d'$ :  $F(2,22)=0.844$ ,  $p > 0.05$ ; reaction time:  $F(2,22)=0.188$ ,  $p > 0.05$ ).
- The number of items stored in VSTM ( $K$ ) increased from set-size 1 to set-size 4 ( $t(11)=10.0$ ,  $p < 0.01$ ), but remained constant (at about 3 items) from set-size 4 to set-size 8 ( $t(11)=0.764$ ,  $p > 0.05$ ).

- Task performance ( $d'$ ) decreased from set-size 1 to set-size 4 ( $t(11)=4.86$ ,  $p<0.01$ ) and again from set-size 4 to set-size 8 ( $t(11)=6.44$ ,  $p<0.01$ ).
- Reaction time (RT) increased from set-size 1 to set-size 4 ( $t(11)=9.94$ ,  $p<0.01$ ) and again from set-size 4 to set-size 8 ( $t(11)=4.69$ ,  $p<0.01$ ).

Performance in the IM task was high (95-99%) and comparable to that reported by Todd and Marois (2004; 96-98%). Performance in the ESA and ESTA tasks was also consistently high (98-99% and 95-97% respectively).

In a repeated measures analysis of variance (ANOVA), with factors of task (VSTM, IM, ESA, ESTA) and set-size (1, 4, 8), performance ( $d'$ ) on the flanking verbal memory test was independent of task ( $F(3,33)=1.02$ ,  $p>0.05$ ) and set-size ( $F(2,22)=1.09$ ,  $p>0.05$ ), with no interaction ( $F(6,66)=0.648$ ,  $p>0.05$ ). RT for the verbal memory response was affected by task ( $F(3,33)=4.02$ ,  $p<0.05$ ), probably reflecting the fact that in the VSTM and ESTA tasks there was a shorter interval between the two required responses leading to a slight increase in the second reaction time. However, as with  $d'$ , reaction time for the verbal memory response was unaffected by set-size ( $F(2,22)=0.442$ ,  $p>0.05$ ) or the interaction of set-size and task ( $F(6,66)=0.448$ ,  $p>0.05$ ).

With respect to the primary response to the visual stimuli, a similar ANOVA with factors of task and set-size showed significant effects of both manipulations on both  $d'$  and RT (all  $p<0.01$ ). The effect of set-size was examined in more detail for each task with one-way ANOVAs and planned  $t$ -tests: In the VSTM task,  $d'$  decreased and RT increased with both increases in set-size as described above. In the IM task,  $d'$  remained constant from set-size 1 to set-size 4 ( $t(11)=1.15$ ,  $p>0.05$ ) but increased from set-size 4 to set-size 8 ( $t(11)=2.24$ ,  $p<0.05$ ), while RT remained constant from set-size 1 to set-size 4 ( $t(11)=0.795$ ,  $p>0.05$ ) and decreased from set-size 4 to set-size 8 ( $t(11)=5.48$ ,  $p<0.01$ ). Neither  $d'$  nor RT were significantly affected by set-size in the ESA task ( $F(2,22)=1.45$ ,  $p>0.05$  and  $F(2,22)=0.163$ ,  $p>0.05$ , respectively) or in the ESTA task ( $F(2,22)=0.976$ ,  $p>0.05$  and  $F(2,22)=0.076$ ,  $p>0.05$ , respectively).

Given that the stimuli in the IM and ESA tasks were identical, it is appropriate to compare performance between them. Such comparisons between the other tasks were not performed because the stimulus differences would complicate their interpretation. RT showed a marginally significant decrease in the ESA task compared to the IM task ( $F(1,11)=4.99$ ,  $p=0.05$ ). The ESA task was found to be slightly, but significantly, easier than the IM task, as measured either by  $d'$  ( $F(1,11)=8.10$ ,  $p<0.05$ ) or percent correct ( $F(1,11)=7.89$ ,  $p<0.05$ ).

### *Neuroimaging*

In a whole-brain analysis, with subjects as a random effect, a contrast of the  $K$ -weighted regressor during the VSTM task revealed significant activity ( $p<0.05$  FDR-corrected) bilaterally along the IPS/IOS and in ventral and lateral occipital cortex, as well as in numerous other regions including the central and pre-central sulci, cerebelli, cingulate cortex extending into the pre-supplementary motor areas, the hippocampi, insulae, inferior frontal regions and subcortical structures (fig. 3A left panel, and fig. 3B, left column). This activity represents the component of the BOLD response which correlates with the number of items stored in VSTM, after accounting for any components which vary linearly with set-size or occur on each trial independently of set-size. A similar, though weaker, pattern of activity was revealed in the occipital and parietal lobes by the same regressor in the ESTA task (fig. 3A right panel, and fig. 3B, right column). No voxels exceeded a whole-brain-corrected threshold for the contrast of the  $K$ -weighted regressor in the IM or ESA tasks, or for the linear set-size-weighted regressor in any task. However, examination of the  $K$ -weighted regression coefficients from the IM task reveals a pattern that shows some similarity to that in the VSTM and ESTA tasks, especially along the IPS (fig. 3B, 2nd column).

In order to probe in more detail the activity in some of these regions, the mean BOLD signal change was extracted from regions of interest (ROIs) centred on the coordinates reported by Todd and Marois (2004) and illustrated in figure 3B. Figure 4A shows the best fitting canonical haemodynamic

response function to the sample display at each set-size, across the three ROIs and four tasks. Since the design was optimised to examine differential responses across set-sizes rather than to estimate the absolute evoked response to a specific set-size, plotted responses are normalised to a set-size of 1 by subtracting the response when only a single item was presented. The absolute signal change at each set-size (provided in table 1) shows a negative offset in some conditions, especially in the occipital ROI. The significance of this is difficult to interpret due to little power to separate responses to the sample display from other sub-processes within each trial. Incorporating temporal jitter between the auditory stimuli, sample display and probe display would allow this to be resolved in future experiments. In figure 4B, the same data are re-plotted showing the peak estimated BOLD signal change increase as a function of set-size. The behavioural measure of the estimated number of objects maintained in the VSTM task ( $K$ ) is included on each axis for comparison. Figure 4C shows the coefficients associated with the regressors modelling the components of the BOLD signal that vary across set-size in a linear, and  $K$ -weighted manner. The data from this analysis is provided in table 2.

Primary interest was in the posterior parietal (IPS/IOS) ROI, where the signal change associated with the  $K$ -weighted regressor was significantly greater than zero in the VSTM, IM and ESTA tasks (all  $p < 0.05$ , 1-tailed; fig. 4C, upper panel). The  $K$ -weighted activity in the IPS was significantly greater during the IM task than the ESA task ( $t(11) = 2.88$ ,  $p < 0.05$ , 2-tailed). Since these tasks were identical in terms of stimuli, the difference can only be attributed to the different task instructions. For the IPS, in no task was the signal change associated with the linear set-size-weighted regressor significantly greater than zero. In the anterior cingulate (ACC) ROI, the only significant effect was for the  $K$ -weighted regressor during the VSTM task (fig. 4C, middle panel). In the ventrolateral occipital (VO) ROI, a significant effect of the  $K$ -weighted regressor was seen during the VSTM and ESTA tasks, along with a significant effect of the linear set-size-weighted regressor during the ESA and ESTA tasks (fig. 4C, lower panel).

In a supplementary analysis, an extra regressor was included in the design matrix to model those memory trials in which the probe disc was taken from the sample colour set, but occurred at a different position. The regressor modelling these events detected significant group activity in both the ACC and IPS ROIs (both  $p < 0.05$ , 2-tailed), but importantly, in all ROIs, its inclusion had a negligible influence on the other regressors.

## Discussion

The main findings of the current study are that (1) activity in the posterior IPS increases with the number of presented objects in simple perceptual tasks as well as in a visual working memory task; (2) this perceptual load-related activity has a similar capacity limit to that observed during the VSTM task, and (3) the observation of load-related activity in non-working memory conditions is sensitive to the particular task demands imposed.

Many aspects of these results are consistent with previous studies, especially those of Todd and Marois (2004), both in terms of the behavioural measures and, importantly, in terms of the replication that activity in the posterior IPS is better explained by a function that asymptotes when the number of display items exceeds working memory capacity, than by a linear function of set-size. However, there are also interesting differences. In the VSTM task, we found significant  $K$ -weighted activity across a wider brain network and in the ventral occipital and anterior cingulate ROIs we found a significant positive effect for the  $K$ -weighted regressor, but not for the linear set-size-weighted regressor. During the IM task, we observed a significant correlation of IPS activity with the  $K$ -weighted regressor, and we failed to find a significant linear effect of set-size in the ventral occipital region.

There are several possible reasons why the present study found parietal  $K$ -weighted activity in perceptual tasks, in contrast to the findings of Todd and Marois (2004). Perhaps apparently perceptual tasks that activate this region involve automatic encoding of object representations into



short-term memory, despite no obvious requirement to do so. This leaves the question of why such automatic encoding would have occurred in the current IM task, but not in that of Todd and Marois. It is conceivable that the current multi-task context in which half the subjects performed the IM task directly after the VSTM task could have encouraged such a task-set. Alternatively, there may have been some sensitivity to perceptual load in Todd and Marois' control experiment, which failed to reach significance due to a lower effect size or reduced power for other task-related or methodological reasons. This would be consistent with the more extensive activations observed during the VSTM task in the current study. Indeed, amongst the additional activated areas found in the whole-brain analysis were those regions identified by Todd and Marois when they relaxed their statistical threshold. Another region strongly activated in the VSTM task was around the junction of the precentral sulcus and the superior frontal sulcus, near to the frontal eye fields (FEF). Signal change in this region has previously been noted to asymptote around set-size two (Culham and others 2001; Todd and Marois 2005), and may reflect planning of upcoming saccades or attention shifts (Corbetta and others 1998). In the current study, the brief, foveal presentation of stimuli was designed to encourage fixation and render saccades unhelpful, however covert attention shifts might constitute an important rehearsal mechanism in working memory (Awh and Jonides 2001; Awh and others 2006; Smyth 1996).

A potential confound in the current VSTM task was that for change trials with a single sample item, the probe item was always a novel colour, however for change trials at the higher set-sizes, the probe had an equal probability of being a novel colour or being a colour used elsewhere in the sample display, possibly causing response conflict. To rule out this explanation for the BOLD set-size effect, the data were re-analysed with the latter events explicitly modelled. While these events produced significant group activity in both the ACC and IPS ROIs, which may reflect an additional element of response conflict, modelling them separately had negligible influence on the other regressors. This suggests that although they increased activity within the ACC and IPS, this is insufficient to explain the activity that correlates with perceptual and working memory load.

The current study failed to find significant positive correlations between the BOLD signal and the linear set-size regressor in either the whole brain analysis or any of the ROIs during the VSTM or IM tasks. It would be inappropriate to draw strong conclusions from such a null result, especially for the occipital ROI, where such an effect was observed during the ESA and ESTA tasks. It remains possible that, with greater statistical power, a linear effect of set-size would have been detected in the occipital ROI during the IM task. The observation that, during the VSTM task, activity in the occipital ROI is better explained by the  $K$ -weighted regressor than the linear set-size weighted regressor is consistent with a recent study (Xu and Chun 2006), although their ROI lay approximately 10mm more lateral than the current occipital ROI. The reduced linear effect of set-size in VO during the VSTM task relative to the perceptual tasks ( $p < 0.05$ , 2-tailed) is also apparent in the plots of signal change against set-size (fig. 4B, lower panel). One might speculate that representations in earlier visual areas reflect an interaction between incoming sensory input and feedback biasing signals from higher levels (e.g. Friston 2005). During perceptual tasks, this balance may favour the visual input, whereas during the memory delay, earlier visual areas may be recruited by feedback connections from PPC, to help maintain the limited number of higher level representations.

It should be noted that the use of relatively short trials in these experiments means that it is not possible here to separate the sluggish haemodynamic response into components reflecting sub-processes of each task, for example encoding, maintenance and retrieval in the VSTM task. Studies which have used slow event-related fMRI designs (Todd and Marois 2004; Xu and Chun 2006), or electrophysiology (Busch and Herrmann 2003), suggest, however, that the activity reflecting object-based capacity limits is likely to correspond to processes occurring during encoding and memory maintenance more than memory retrieval.

*What aspects of a perceptual task might determine whether it evokes a capacity limited response in PPC?*

Although we have demonstrated that the IPS has a capacity-limited response to perceptual load as well as to working memory load, there was a noticeable lack of such an effect in the ESA task, despite its being visually identical to the IM task. This sensitivity of IPS perceptual load-related activity to the specific perceptual task could potentially be explained by a number of factors, which include the extent of spatial attention, the duration of attention, the dependence of the task on spatial information, task difficulty and the binding of properties to spatiotemporal object tokens. These will be discussed in turn.

The hypothesis motivating this study was that, even in the absence of an explicit working memory requirement, as attentional demands were increased in the perceptual tasks, significant set-size-related BOLD signal changes would be observed in posterior parietal cortex. The IM and ESA tasks were designed to be physically identical, but with the required spatial scope of attention being increased in the latter. We therefore expected to see more reliable  $K$ -weighted activity in the ESA task if the spatial extent of attention were important (at least at the scale employed in the current manipulation), but no difference if it were not. The reverse finding that the  $K$ -weighted activity in the IPS ROI was significantly greater during the IM task than the ESA task therefore suggests that perceptual load-related activity is not dependent on increasing the extent of spatial attention at this scale.

It was also considered that temporally sustained attention to the stimulus representations, whether in working memory or constantly visible, may be required for distinct object representations to build up in posterior parietal cortex, and so set-size-weighted BOLD activity would not be measured if attention were disengaged from the stimuli too quickly. That sustained attention is required for segregated object representations to build up is suggested from studies of visual apparent motion (Anstis and others 1985) and auditory streaming (Cusack and others 2004), where the percept of

multiple, separable objects/streams increases in likelihood as attention to an ambiguous stimulus is maintained. This motivated the design of the ESTA task, where attention to the sample display was prolonged. If perceptual load-related activity were dependent on attentional duration, we expected to see a greater set-size effect in the VSTM and ESTA tasks, where the items (or their representations in memory) were being attended for over a second, compared to the IM and ESA tasks, where a response was made immediately and attention could thus be disengaged much earlier. Given that the stimuli in the latter two tasks were identical, the significant difference in *K*-weighted activity between the IM and ESA tasks makes this hypothesis unlikely, although it should be noted that the mean reaction time for the ESA task was slightly shorter than for the IM task.

Considering the role of PPC in the processing of spatial information (e.g. Owen and others 1996; Schluppeck and others 2006; Sereno and others 2001; Silver and others 2005), and the fact that different objects typically occupy unique spatial locations, one might surmise that the observed load sensitivity corresponds specifically to the processing of object locations rather than object representations more generally. In this case it should be seen in those conditions where location is relevant to the task. Although this would explain the observed pattern in the VSTM, IM and ESA tasks (with the ESA task being the only one where location was irrelevant and the only one where IPS activity did not vary with set-size), it would predict no significant set-size effect in the ESTA task where position was equally irrelevant. This hypothesis was therefore also rejected. That distinct spatial locations are not necessary to observe an effect of object number on IPS activity is also suggested by paradigms using sequential presentation at a single location (Linden and others 2003; Xu and Chun 2006) as well as experiments from our lab using simultaneous presentation, where ‘objects’ were either transparent overlapping dot surfaces or auditory streams coming from the same part of space (Cusack 2005).

Along with the marginally significant difference in reaction times between the IM and ESA tasks, the ESA task was found to be slightly, but significantly, easier than the IM task, as measured either

by  $d'$  or percent correct. So, although the  $K$ -weighted activity does not seem to index task difficulty directly (difficulty increased linearly with set-size in the VSTM task and did not increase with set-size in any of the perceptual tasks), it remains possible that this activity is modulated by the overall difficulty of the task, becoming more pronounced as the task becomes more challenging.

In retrospect, it seems plausible that although the whole stimulus array had to be assessed in the ESA task, because only a single feature (redness) had to be detected, this could have been accomplished in an efficient, parallel manner, without needing to bind this colour property to an individual object. In contrast, in the IM task, the central disc must be bound to its location, which may require the items to be individuated rather than processed as a group<sup>1</sup>. Following this logic, it may be that detecting the staggered offset in the ESTA task requires binding of each object to its time of offset. Along similar lines, segregating the discs may be required to enhance the signal to noise ratio by binding properties to individual items rather than binding a mean property to the group as a whole. This would be consistent with classical models of visual search, where search for targets defined by features coded automatically in early vision (e.g. redness in the IM task) can occur in parallel across the array, whereas search for the absence of a primitive feature (e.g. contrast in the ESTA task) benefits from segmenting the array into individual objects (Treisman and Gormican 1988). There is also evidence that elements can be selectively grouped into higher-order objects depending on the requirements of a visual search task (Duncan and Humphreys 1989; Treisman 1982) or a multiple object tracking task (Yantis 1992). It is therefore possible that the observation of perceptual-load related activity in the IPS is dependent on flexible grouping/segmentation of hierarchical object representations and that the optimal degree of segmentation depends on the salience of the discrimination required or the necessity to bind the discriminanda to particular spatiotemporal tokens. This idea may not be unrelated to the explanation in terms of task difficulty i.e. as the

---

<sup>1</sup> Of course, it may seem optimal in the IM task to just encode the central position, while completely ignoring any surrounding discs. The fact that subjects did not do this, however, is indicated by the significant difference in performance and reaction time in this task between set-sizes of four and eight, which were identical as far as the central array position was concerned. It is worth noting that for set-size 8 and some configurations of four items, although

required discriminations become more demanding, task difficulty increases and a greater level of segmentation may be required to compensate. Further work will be required to disentangle these possibilities.

The current observation that posterior parietal activity increases with the number of objects that are presented is consistent with previous studies in which storage of objects in working memory has not been emphasised (Cusack 2005; Fang and He 2005; Hon and others 2006; Jovicich and others 2001; Ritzl and others 2003; Tomasi and others 2004). That this activity asymptotes by the time the number of objects has risen to around four is also consistent with neuroimaging studies of multiple object tracking (Culham and others 2001; Jovicich and others 2001; Tomasi and others 2004), a paradigm in which the capacity limit is often thought to reflect pre-attentive or attentional selection processes rather than working memory processes (Pylyshyn Z. W. and Storm 1988). The visual indexing mechanism suggested by Pylyshyn and colleagues to account for performance in MOT and other tasks that show evidence of limited capacity parallel processes in early vision (e.g. subitizing and the detection of illusory line motion; Pylyshyn Z. 1989; Schmidt and others 1998; Trick and Pylyshyn 1994) is proposed to consist of a limited number of demonstrative pointers which are linked to objects prior to any assignment of properties. They serve to individuate the objects, to provide rapid access to their properties, and to maintain the identities of spatiotemporally continuous entities despite changes in spatial location, sensory properties or classification label. They may be further elaborated into temporary object-files (Kahneman and others 1992). It is possible that the capacity-limited posterior parietal activity observed here in simple perceptual tasks reflects this process of individuating a small number of relevant objects, so that properties such as colour can then be bound to them. This would also be consistent with a recent study which demonstrated that a similar region was better activated by discrete visual stimuli than by continuous, analogue stimuli, from which it is presumably more difficult to individuate distinct items (Castelli and others 2006).

---

processing of the surrounding items is not *necessary* to complete the task, it is *sufficient*, because the global configuration of the array determines whether an item is present in the central position.

## Conclusion

This study demonstrates that the region of the IPS where BOLD activity reflects the limited number of objects held in working memory also shows a similar capacity-limited response under certain perceptual conditions that involve no explicit memory requirements. This would be consistent with proposals that the IPS activity seen during VSTM tasks is a product of attentional selection processes rather than representing the contents of working memory *per se* (Cowan 2001; Majerus and others 2007; Marois and Ivanoff 2005). Furthermore, this effect is sensitive to the particular perceptual task imposed, even when the stimuli are identical, perhaps reflecting a task requirement for the binding of properties to spatiotemporally individuated object tokens. Thus, the physiological markers of the number of internal object representations developed by Todd and Marois (2004) and by Vogel and Machizawa (2004) provide a useful tool to investigate how stimulus properties and task demands constrain the visual system's individuation of objects from a scene in perception and working memory. Given the highly limited capacity of visual short-term memory and of visuospatial attention, the ability to manage these resources flexibly is crucial. Dynamic grouping and segregation of perceptual object representations according to current task demands may be one mechanism of achieving such flexibility, and this computation may occur within posterior parietal cortex.

## Notes

DJM was supported by the Medical Research Council (UK). We thank John Duncan for helpful comments on the manuscript. *Conflict of interest*: None declared.

Address correspondence to Daniel J. Mitchell, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, CB2 7EF, UK. Email: [daniel.mitchell@mrc-cbu.cam.ac.uk](mailto:daniel.mitchell@mrc-cbu.cam.ac.uk)

## References

- Alvarez GA, Cavanagh P. 2004. The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychol Sci* 15: 106-111.
- Anstis S, Giaschi D, Cogan AI. 1985. Adaptation to apparent motion. *Vision Res* 25: 1051-1062.
- Awh E, Jonides J. 2001. Overlapping mechanisms of attention and spatial working memory. *Trends Cogn Sci* 5: 119-126.
- Awh E, Vogel EK, Oh SH. 2006. Interactions between attention and working memory. *Neuroscience* 139: 201-208.
- Baddeley A 1986 *Working memory*. New York: Oxford University Press.
- Bor D, Duncan J, Wiseman RJ, Owen AM. 2003. Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron* 37: 361-367.
- Busch NA, Herrmann CS. 2003. Object-load and feature-load modulate EEG in a short-term memory task. *Neuroreport* 14: 1721-1724.
- Castelli F, Glaser DE, Butterworth B. 2006. Discrete and analogue quantity processing in the parietal lobe: a functional MRI study. *Proc Natl Acad Sci U S A* 103: 4693-4698.
- Corbetta M, Akbudak E, Conturo TE, Snyder AZ, Ollinger JM, Drury HA, Linenweber MR, Petersen SE, Raichle ME, Van Essen DC, Shulman GL. 1998. A common network of functional areas for attention and eye movements. *Neuron* 21: 761-773.
- 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.
- Culham JC, Cavanagh P, Kanwisher NG. 2001. Attention response functions: characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron* 32: 737-745.
- Cusack R, Deeks J, Aikman G, Carlyon RP. 2004. Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J Exp Psychol Hum Percept Perform* 30: 643-656.
- Cusack R. 2005. The intraparietal sulcus and perceptual organization. *J Cogn Neurosci* 17: 641-651.
- Davis G, Holmes A. 2005. The capacity of visual short-term memory is not a fixed number of objects. *Mem Cognit* 33: 185-195.
- Duncan J. 1984. Selective attention and the organization of visual information. *J Exp Psychol Gen* 113: 501-517.
- Duncan J, Humphreys GW. 1989. Visual search and stimulus similarity. *Psychol Rev* 96: 433-458.
- Fang F, He S. 2005. Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat Neurosci* 8: 1380-1385.
- Ferreira CT, Verin M, Pillon B, Levy R, Dubois B, Agid Y. 1998. Spatio-temporal working memory and frontal lesions in man. *Cortex* 34: 83-98.



- Friston K. 2005. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360: 815-836.
- Greenwood JA, Edwards M. 2006. Pushing the limits of transparent-motion detection with binocular disparity. *Vision Res* 46: 2615-2624.
- Hikosaka O, Miyauchi S, Shimojo S. 1993. Focal visual attention produces illusory temporal order and motion sensation. *Vision Res* 33: 1219-1240.
- Hon N, Epstein RA, Owen AM, Duncan J. 2006. Frontoparietal activity with minimal decision and control. *J Neurosci* 26: 9805-9809.
- Humphreys GW. 1998. Neural representation of objects in space: a dual coding account. *Philos Trans R Soc Lond B Biol Sci* 353: 1341-1351.
- Jiang Y, Olson IR, Chun MM. 2000. Organization of visual short-term memory. *J Exp Psychol Learn Mem Cogn* 26: 683-702.
- Jiang Y, Chun MM, Olson IR. 2004. Perceptual grouping in change detection. *Percept Psychophys* 66: 446-453.
- Jovicich J, Peters RJ, Koch C, Braun J, Chang L, Ernst T. 2001. Brain areas specific for attentional load in a motion-tracking task. *J Cogn Neurosci* 13: 1048-1058.
- Kahneman D, Treisman A, Gibbs BJ. 1992. The reviewing of object files: object-specific integration of information. *Cognit Psychol* 24: 175-219.
- Kaufman EL, Lord MW. 1949. The discrimination of visual number. *Am J Psychol* 62: 498-525.
- Linden DE, Bittner RA, Muckli L, Waltz JA, Kriegeskorte N, Goebel R, Singer W, Munk MH. 2003. Cortical capacity constraints for visual working memory: dissociation of fMRI load effects in a fronto-parietal network. *Neuroimage* 20: 1518-1530.
- Luck SJ, Vogel EK. 1997. The capacity of visual working memory for features and conjunctions. *Nature* 390: 279-281.
- Majerus S, Bastin C, Poncelet M, Van der Linden M, Salmon E, Collette F, Maquet P. 2007. Short-term memory and the left intraparietal sulcus: focus of attention? Further evidence from a face short-term memory paradigm. *Neuroimage* 35: 353-367.
- Marois R, Ivanoff J. 2005. Capacity limits of information processing in the brain. *Trends Cogn Sci* 9: 296-305.
- O'Craven KM, Downing PE, Kanwisher N. 1999. fMRI evidence for objects as the units of attentional selection. *Nature* 401: 584-587.
- Olson IR, Jiang Y. 2002. Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Percept Psychophys* 64: 1055-1067.

- Owen AM, Milner B, Petrides M, Evans AC. 1996. Memory for object features versus memory for object location: a positron-emission tomography study of encoding and retrieval processes. *Proc Natl Acad Sci U S A* 93: 9212-9217.
- Owen AM. 2004. Working memory: imaging the magic number four. *Curr Biol* 14: R573-574.
- Pashler H. 1988. Familiarity and visual change detection. *Percept Psychophys* 44: 369-378.
- Paul MA, Tipper SP. 2003. Object-based representations facilitate memory for inhibitory processes. *Exp Brain Res* 148: 283-289.
- Phillips WA. 1974. On the distinction between sensory storage and short-term visual memory. *Percept Psychophys* 16: 283-290.
- Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S. 2004. Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron* 44: 547-555.
- Pochon JB, Levy R, Poline JB, Crozier S, Lehericy S, Pillon B, Deweer B, Le Bihan D, Dubois B. 2001. The role of dorsolateral prefrontal cortex in the preparation of forthcoming actions: an fMRI study. *Cereb Cortex* 11: 260-266.
- Postle BR. 2006. Working memory as an emergent property of the mind and brain. *Neuroscience* 139: 23-38.
- Pylyshyn Z. 1989. The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. *Cognition* 32: 65-97.
- Pylyshyn ZW, Storm RW. 1988. Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spat Vis* 3: 179-197.
- Pylyshyn ZW. 2001. Visual indexes, preconceptual objects, and situated vision. *Cognition* 80: 127-158.
- Ritzl A, Marshall JC, Weiss PH, Zafiris O, Shah NJ, Zilles K, Fink GR. 2003. Functional anatomy and differential time courses of neural processing for explicit, inferred, and illusory contours. An event-related fMRI study. *Neuroimage* 19: 1567-1577.
- Schluppeck D, Curtis CE, Glimcher PW, Heeger DJ. 2006. Sustained activity in topographic areas of human posterior parietal cortex during memory-guided saccades. *J Neurosci* 26: 5098-5108.
- Schmidt WC, Fisher BD, Pylyshyn ZW. 1998. Multiple-location access in vision: evidence from illusory line motion. *J Exp Psychol Hum Percept Perform* 24: 505-525.
- Scholl BJ, Pylyshyn ZW, Feldman J. 2001. What is a visual object? Evidence from target merging in multiple object tracking. *Cognition* 80: 159-177.
- Sereno MI, Pitzalis S, Martinez A. 2001. Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science* 294: 1350-1354.

- Silver MA, Ress D, Heeger DJ. 2005. Topographic maps of visual spatial attention in human parietal cortex. *J Neurophysiol* 94: 1358-1371.
- Smyth MM. 1996. Interference with rehearsal in spatial working memory in the absence of eye movements. *Q J Exp Psychol A* 49: 940-949.
- Starkey P, Cooper RG, Jr. 1980. Perception of numbers by human infants. *Science* 210: 1033-1035.
- Tipper SP, Behrmann M. 1996. Object-centered not scene-based visual neglect. *J Exp Psychol Hum Percept Perform* 22: 1261-1278.
- Todd JJ, Marois R. 2004. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428: 751-754.
- Todd JJ, Marois R. 2005. Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cogn Affect Behav Neurosci* 5: 144-155.
- Tomasi D, Ernst T, Caparelli EC, Chang L. 2004. Practice-induced changes of brain function during visual attention: a parametric fMRI study at 4 Tesla. *Neuroimage* 23: 1414-1421.
- Treisman A. 1982. Perceptual grouping and attention in visual search for features and for objects. *J Exp Psychol Hum Percept Perform* 8: 194-214.
- Treisman A, Gormican S. 1988. Feature analysis in early vision: evidence from search asymmetries. *Psychol Rev* 95: 15-48.
- Trick LM, Pylyshyn ZW. 1994. Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol Rev* 101: 80-102.
- Vergilino-Perez D, Findlay JM. 2006. Between-object and within-object saccade programming in a visual search task. *Vision Res* 46: 2204-2216.
- Viswanathan L, Mingolla E. 2002. Dynamics of attention in depth: evidence from multi-element tracking. *Perception* 31: 1415-1437.
- Vogel EK, Machizawa MG. 2004. Neural activity predicts individual differences in visual working memory capacity. *Nature* 428: 748-751.
- Vogel EK, Woodman GF, Luck SJ. 2006. The time course of consolidation in visual working memory. *J Exp Psychol Hum Percept Perform* 32: 1436-1451.
- Wilken P, Ma WJ. 2004. A detection theory account of change detection. *J Vis* 4: 1120-1135.
- Xu Y, Chun MM. 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440: 91-95.
- Yantis S. 1992. Multielement visual tracking: attention and perceptual organization. *Cognit Psychol* 24: 295-340.

## Tables

**Table 1: Results of the ROI in which the *K*-weighted and linear-set-size-weighted components of the BOLD response were modelled. Coordinates given in Talairach space.**

ROI	Task	Regressor	Mean regression	<i>t</i>	<i>p</i>
			coefficient	(df=11)	(1-tailed)
IPS Left: -22, -65, 42 Right: 23, -59, 45	VSTM	<i>K</i> -weighted	0.939	4.472	<0.001*
		Set-size-weighted	-0.317	-2.957	0.993
	IM	<i>K</i> -weighted	0.619	2.493	0.015*
		Set-size-weighted	-0.240	-1.048	0.841
	ESA	<i>K</i> -weighted	-0.062	-0.555	0.705
		Set-size-weighted	0.121	0.870	0.202
	ESTA	<i>K</i> -weighted	0.477	3.492	0.003*
		Set-size-weighted	-0.056	-0.567	0.709
	VSTM	<i>K</i> -weighted	0.372	3.130	0.005*
		Set-size-weighted	0.011	0.167	0.435
ACC Right: 8, 19, 30	IM	<i>K</i> -weighted	0.058	0.332	0.373
		Set-size-weighted	0.013	0.082	0.468
	ESA	<i>K</i> -weighted	-0.044	-0.365	0.639
		Set-size-weighted	0.011	0.123	0.452
	ESTA	<i>K</i> -weighted	0.130	1.242	0.120
		Set-size-weighted	-0.011	-0.139	0.554
VO Left: -34, -68, 0 Right: 32, -77, -9	VSTM	<i>K</i> -weighted	0.410	2.732	0.010*
		Set-size-weighted	-0.067	-0.591	0.717
	IM	<i>K</i> -weighted	0.081	0.488	0.318
		Set-size-weighted	0.168	1.054	0.157
	ESA	<i>K</i> -weighted	-0.053	-0.571	0.710
		Set-size-weighted	0.218	2.290	0.021*
	ESTA	<i>K</i> -weighted	0.341	4.225	0.001*
		Set-size-weighted	0.144	2.772	0.009*

**Table 2: Results of the ROI analysis in which each set-size was modelled separately. Coordinates given in Talairach space.**

ROI	Task	Regressor	% signal change (from ROI session mean)	Increase in % signal change relative to set-size 1	t-test of increase from lower set-size	
					t	p
					(df=11)	(1-tailed)
IPS  Left: -22, -65, 42  Right: 23, -59, 45	VSTM	Set-size 1	0.0376			
		Set-size 4	0.1928	0.1553	4.6255	0.0004*
		Set-size 8	0.1675	0.1299	-1.9965	0.9644
	IM	Set-size 1	-0.0197			
		Set-size 4	0.0752	0.0949	3.4092	0.0029*
		Set-size 8	0.0531	0.0728	-0.6975	0.7500
	ESA	Set-size 1	0.0140			
		Set-size 4	0.0150	0.0010	0.0725	0.4718
		Set-size 8	0.0313	0.0173	0.7435	0.2364
	ESTA	Set-size 1	-0.0397			
		Set-size 4	0.0489	0.0885	3.8461	0.0014*
		Set-size 8	0.0482	0.0879	-0.0505	0.5197
ACC  Right: 8, 19, 30	VSTM	Set-size 1	-0.0588			
		Set-size 4	0.0376	0.0964	4.4071	0.0005*
		Set-size 8	0.0503	0.1091	1.3022	0.1097
	IM	Set-size 1	-0.0459			
		Set-size 4	-0.0260	0.0198	0.8620	0.2035
		Set-size 8	-0.0198	0.0260	0.2439	0.4059
	ESA	Set-size 1	-0.0120			
		Set-size 4	-0.0228	-0.0108	-0.5536	0.7045
		Set-size 8	-0.0211	-0.0092	0.1190	0.4537
	ESTA	Set-size 1	-0.0609			
		Set-size 4	-0.0286	0.0322	1.7412	0.0547
		Set-size 8	-0.0315	0.0294	-0.2032	0.5787
VO  Left: -34, -68, 0  Right: 32, -77, -9	VSTM	Set-size 1	-0.2464			
		Set-size 4	-0.1462	0.1002	4.0566	0.0009*
		Set-size 8	-0.1469	0.0995	-0.0334	0.5130
	IM	Set-size 1	-0.2094			

	Set-size 4	-0.1605	0.0489	2.1334	0.0281*
	Set-size 8	-0.1265	0.0829	1.1439	0.1385
ESA	Set-size 1	-0.1838			
	Set-size 4	-0.1641	0.0197	1.6110	0.0677
	Set-size 8	-0.1199	0.0639	2.4092	0.0173*
ESTA	Set-size 1	-0.2290			
	Set-size 4	-0.1127	0.1163	6.0443	<0.0001*
	Set-size 8	-0.0714	0.1576	3.7351	0.0016*

## Captions to figures

Figure 1: Schematic of the stimulus presentation sequence. Each trial began with a visual ‘ready’ prompt, then two auditory digits followed by a fixation point on a mid-grey background. A sample display of 1, 4 or 8 differently coloured discs appeared for 150ms in the VSTM, IM and ESA tasks and for 750-1350ms in the ESTA task. In the VSTM task the sample was followed by a 1200ms retention interval, probe disc and response interval, while in the other tasks the response interval followed directly from the offset of the sample display. All trials ended with two visually presented digits, a second response and a variable duration inter-trial interval.

Figure 2: Mean behavioural performance across the group. In the upper panel, the number of items-worth of information encoded in short-term memory ( $K$ ) is plotted as a function of set-size in the VSTM task. Performance ( $d'$ ) and reaction time (ms) are plotted against set-size for all tasks, in the middle and lower panels respectively. Error bars indicate  $\pm 1$  standard error of the mean across subjects.

Figure 3: fMRI whole-brain group results. (A) t-statistics representing  $K$ -weighted BOLD activity during the VSTM task (left) and the ESTA task (right). Illustrated activations are significant at  $p < 0.05$  after whole-brain FDR correction, and are rendered onto the Colin template brain. The IM and ESA tasks revealed no supra-threshold voxels. Dashed blue lines indicate the positions of the cross-sections shown below. (B) Regression coefficients for the regressor modelling  $K$ -weighted BOLD signal changes, overlaid on selected axial slices. Z-coordinates are given in MNI space; black, dotted contours indicate the significant activations rendered in panel A; white, dotted contours indicate ROIs centred on coordinates taken from Todd & Marois (2004).

Figure 4: BOLD activity extracted from the ROIs illustrated in figure 3B. The anterior cingulate (ACC) ROI was a single locus, while the posterior parietal (IPS) and ventral occipital (VO) ROIs pooled data across both hemispheres. (A) Estimated BOLD time-courses following onset of the sample display at each set-size. Event-related amplitude fits of the canonical haemodynamic response function are normalised relative to set-size 1 within each task and ROI. Dotted lines indicate  $\pm 1$  standard error. (B) Peak, normalised, percent signal change, plotted as a function of set-size. The estimated number of objects maintained in the VSTM task ( $K$ ) is plotted on each axis for comparison. (C) Blue bars represent mean regression coefficients for the regressor modelling  $K$ -weighted BOLD signal changes, while red bars represent coefficients for the regressor modelling a linear effect of set-size. Different task conditions are presented along the y-axes for each ROI. Error bars indicate  $\pm 1$  standard error of the mean across subjects.