DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

The Neural Fate of Task-Irrelevant Features in Object-Based Processing

The Harvard community has made this article openly available.

Please share how this access benefits you. Your story matters.

(Article begins on next page)

Behavioral/Systems/Cognitive

The Neural Fate of Task-Irrelevant Features in Object-Based Processing

Yaoda Xu

Department of Psychology, Harvard University, Cambridge, Massachusetts 02138

Objects are one of the most fundamental units in visual attentional selection and information processing. Studies have shown that, during object-based processing, all features of an attended object may be encoded together, even when these features are task irrelevant. Some recent studies, however, have failed to find this effect. What determines when object-based processing may or may not occur? In three experiments, observers were asked to encode object colors and the processing of task-irrelevant object shapes was evaluated by measuring functional magnetic resonance imaging responses from a brain area involved in shape representation. Whereas object-based task-irrelevant shape processing was present at low color-encoding load, it was attenuated or even suppressed at high color-encoding load. Moreover, such object-based processing was short-lived and was not sustained over a long delay period. Object-based processing for task-irrelevant features of attended objects thus does exist, as reported previously; but it is transient and its magnitude is determined by the encoding demand of the task-relevant feature.

Introduction

Decades of vision research has shown that objects are discrete units of visual attentional selection and information processing (Duncan, 1984; Luck and Vogel, 1997; Scholl, 2001). In particular, there exists an object-based processing benefit such that, when attention is directed to one object feature, all other features of that object may be registered, even when these unattended features are task irrelevant. For example, O'Craven et al. (1999) observed higher functional magnetic resonance imaging (fMRI) response unique to feature X when attention was directed to feature Y of the same object than when feature X was located on a completely unattended object. To explain this benefit, researchers argued that different object features can be processed and retained independently and that the object benefit results from features not competing with each other for the same processing resources (Allport, 1971; Wheeler and Treisman, 2002; Xu, 2002a). As such, a task-irrelevant feature may be registered for free when attention is directed to another feature of the same object.

Meanwhile, our visual system is capacity limited and can only process a small amount of visual information at any given time. Given this constraint, it would be highly advantageous for our visual system to only process what is most relevant to the current task and goal of the observer. Indeed, two recent studies found that a task-irrelevant object feature might not be processed when observers were asked to encode another feature from the same object. Woodman and Vogel (2008) found that the rate of information consolidation (Vogel et al., 2006) as well as the event-

related potential (ERP) response amplitude signaling the amount of information retained (Vogel and Machizawa, 2004) differed when observers encoded and retained just the color, just the shape, or the conjunction of these features from each object. Serences et al. (2009) found that fMRI response pattern in the primary visual area only carried information about the task-relevant, but not the task-irrelevant, feature. These studies show that object-based processing is not automatic and observers seem to be able to exert control over which object features are encoded and retained. This is consistent with feature-based information processing (Maunsell and Treue, 2006).

Thus, although some previous studies have left the impression that object-based processing is omnipresent, others have failed to find this effect. How do we reconcile these seemingly conflicting results? More importantly, under what conditions does object-based processing occur? How does the encoding demand of a task-relevant feature affect the encoding of a task-irrelevant feature from the same object? To address these questions, the present study used fMRI and examined the processing of a task-irrelevant feature while manipulating the encoding demand of anther feature from the same object.

Materials and Methods

Participants

Eight paid observers participated in each of the three experiments. There were five, three, and three females in experiments 1, 2, and 3, respectively. Three observers (one female) participated in all three experiments, one observer (female) participated in both experiments 1 and 2, and one observer (female) participated in both experiments 1 and 3. Observers were recruited from the Yale University community with informed consent. They were aged between 18 and 35, right-handed, and had normal or corrected to normal visual acuity and normal color vision. The study was approved by the Yale University School of Medicine Human Investigation Committee.

An additional observer (male) was tested in experiment 1 and another one (female) was tested in experiment 2. Data from these two observers

were excluded from the analyses because of extremely low behavioral performance at display set size 6 (they retained <0.5 color of the six colors presented). These observers most likely gave up entirely on the set size 6 displays, making their fMRI data difficult to interpret. An additional observer (male) was tested in experiment 3. However, this observer reported that he was extremely sleepy during the scan. Consistent with this self-report, this observer retained <0.5 color when only one color was present in the display (a rather easy condition). As such, his data were excluded from the analysis.

Design and procedure

Experiment 1. In each trial, observers viewed a sample display containing one, two, or six colored shapes around the central fixation. After a brief blank delay, they judged whether the centrally presented color probe in the test display matched one of the sample colors (see Fig. 2*A*). Verbal suppression was not used in this experiment because with a visual short-term memory (VSTM) retention interval this brief, verbal recoding of the visual stimulus is rather unlikely (Vogel et al., 2001).

Seven distinctive object shapes, as in the study by Xu and Chun (2006), experiment 4, and seven distinctive colors (red, green, blue, orange, cyan, yellow, white) as in the study by Xu (2007) were used (see Fig. 2*A*). All displays subtended 13.7 × 13.7° of visual angle and were presented on a light gray background. A given object subtended maximally 3.1 × 3.1° of visual angle. Each trial lasted 6 s and consisted of fixation (1000 ms), sample display (200 ms), blank delay (1000 ms), test display/response period (2500 ms), and response feedback (1300 ms) with either a happy face (correct response) or a sad face (incorrect response) at fixation. There were also blank fixation trials in which only a fixation dot was present throughout the 6 s trial duration. Trial presentation order was pseudorandom and balanced in a run (Todd and Marois, 2004; Xu and Chun, 2006). Each observer was tested with two runs, each containing 15 trials for each display condition and lasting 7 min and 9 s.

Experiment 2. This experiment used the same stimuli and design as those of experiment 1, except that observers retained the colored shapes for an extended memory delay period. To prevent verbal recoding of the visual stimuli over the long delay period, observers covertly rehearsed four digits repeatedly throughout each trial (Luck and Vogel, 1997; Todd and Marois, 2004; Xu and Chun, 2006). Following the study by Xu and Chun (2006), experiment 4, a given trial lasted 18 s and consisted of blank fixation (1000 ms), sequential presentation of four different blue digits at the center (250 ms/each), blank fixation (2500 ms), sample display (200 ms), delay period (8300 ms), color probe and response period (2000 ms), color response feedback (500 ms), test digits (presented simultaneously in green color) and response period (2000 ms), and digit response feedback (500 ms). There were also blank fixation trials in which only the digit task, but not the color VSTM task, was present during the trial. As in experiment 1, the presentation order of the different trial types was pseudorandom and balanced in a run. Each observer was tested with six runs, each containing four trials per display set size and lasting 5 min and 30 s. Other aspects of this experiment were identical with that of experiment 1.

Experiment 3. This experiment used the same stimuli and a similar design as that of experiment 1. Display set size was fixed to six objects. A given display can contain either six identical shapes in the same color, six identical shapes in six different colors, six different shapes in the same color, or six different shapes in six different colors (see Fig. 4*A*). After a brief blank delay, observers judged whether the centrally presented color probe in the test display matched one of the sample colors. Each observer was tested with two runs, each containing 15 trials for each display condition and lasting 8 min and 15 s. Other aspects of this experiment were identical with that of experiment 1.

Localizer scans. To define the superior intraparietal sulcus (IPS) region of interest (ROI), a VSTM color experiment was conducted as in the study by Xu and Chun (2006) (Xu, 2007, 2009). A given sample display contained either two to four or six colored squares around the center fixation and appeared for 200 ms. After a 1000 ms blank delay, a test probe appeared in a location occupied by one of the colors in the sample display. The probe either matched the color of the item at the same location in the sample display for no-change trials, or had a color present elsewhere in the sample displays for change trials. Maximum VSTM

capacity for this color–location conjunction task was ~3.5. Seven colors (red, green, blue, cyan, yellow, white, and magenta) were used. As in the main experiments, all displays subtended 13.7 × 13.7° and were presented on a light gray background. Each colored square subtended 2.0 × 2.0°. Each observer was tested with two runs, each containing 12 trials for each display set size and lasting 5 min and 12 s.

To define the lateral occipital complex (LOC) ROI, observers viewed blocks of black object images and blocks of noise images as in the study by Xu and Chun (2006) (Xu, 2007, 2009). Each object image contained six black shapes, created by the same computer algorithm used to generate the displays in the main experiments. This ensured that only LOC regions involved in processing the type of visual objects used in the VSTM experiments were localized. Each image was presented for 500 ms, followed by a 300 ms blank interval before the next image appeared. To ensure attention to the displays, observers fixated at the center and detected a slight spatial jitter, occurring randomly in 1 of every 10 images. Each observer was tested with two runs, each containing 160 black object images and 160 noise images. Each run lasted 4 min and 40 s. Displays used in this localizer scan had the same spatial extent as those in the main experiment.

Because the fMRI data analysis software used could accurately align brain images acquired from the same observers in different scan sessions, localizer scans were not always acquired in the same scanning session with the main experimental scans. This usually happened when the same localizer scans had already been acquired from an observer in a different experiment.

fMRI methods

Observers lay on their back inside a Siemens Trio 3T scanner and viewed the backprojected liquid crystal display with a mirror mounted inside the headcoil. Stimulus presentation and behavioral response collection were controlled by an Apple Powerbook G4 running Matlab with Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). Standard protocols were followed to acquire the anatomical images. A gradient echo pulse sequence (echo time, 25 ms; flip angle, 90°; matrix, 64 × 64) was used for both the main experiments and the localizer scans, with repetition times of 1.5 and 2.0 s, respectively, for the blocked and the event-related runs. Twenty-four 5-mm-thick (3.75 × 3.75 mm in-plane, 0 mm skip) axial slices parallel to the anterior commissure–posterior commissure line were collected.

Data analysis

fMRI data collected were analyzed using BrainVoyager QX. Data preprocessing included slice acquisition time correction, three-dimensional motion correction, linear trend removal, and Talairach space transformation (Talairach and Tournoux, 1988).

A multiple regression analysis was performed separately on each observer on the data acquired from the color VSTM localizer run. The regression coefficient for each set size was weighted by the corresponding behavioral K estimate from that observer for that set size (Todd and Marois, 2004). The superior IPS ROI was defined as voxels that showed a significant activation in the regression analysis (false discovery rate, $q <$ 0.05) and whose Talairach coordinates matched those reported previously (Todd and Marois, 2004). As in our previous study (Xu and Chun, 2006), the LOC ROI was defined as regions in the ventral and lateral occipital cortex whose activations were higher for the black objects than for the noise images (false discovery rate, $q < 0.05$).

These ROIs were overlaid onto the data from the main VSTM experiments and time courses from each observer were extracted. As in previous studies (Kourtzi and Kanwisher, 2000; Todd and Marois, 2004; Xu and Chun, 2006), these time courses were converted to percentage signal change for each stimulus condition by subtracting the corresponding value for the fixation trials and then dividing by that value. In experiments 1 and 3, following previous convention (Todd and Marois, 2004; Xu and Chun, 2006), peak responses were derived by collapsing the time courses for all the conditions and determining the time point of greatest signal amplitude in the averaged response. This was done separately for each observer in each ROI. The resulting peak responses were then averaged across observers.
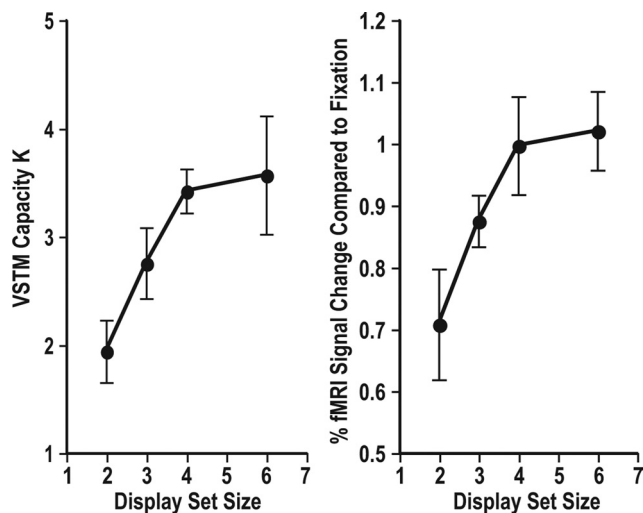
Figure 1. Results of the superior IPS localizer scan from experiment 1. Left, Behavioral performance of the localizer task, as measured by Cowan's K (Cowan, 2001). Right, Peak fMRI responses of this localizer from the superior IPS ROI that it localized. The response patterns were identical between the behavioral and the fMRI measures, with both showing a response increase with increasing display set size and a response plateau at about set size 4. These results are not surprising given how the superior IPS ROI was localized. Nonetheless, they confirm and replicate previous findings and show that the superior IPS localizer was successful in localizing the superior IPS ROI reported in previous studies (Todd and Marois, 2004; Xu and Chun, 2006; Xu, 2007). Errors shown are within-subject SEs.

To confirm that the superior IPS localizer indeed successfully localized the parietal regions that track the maximal number of items an observer can encode and retain from a visual display (Todd and Marois, 2004; Xu and Chun, 2006; Xu, 2007), Figure 1 shows the results of the superior IPS localizer scan from experiment 1. Figure 1, left, plots the behavioral performance of the localizer task, as measured by Cowan's K (Cowan, 2001), whereas Figure 1, right, plots peak fMRI responses of this localizer from the superior IPS ROI that it localized. Because a fixation condition was not included in this localizer scan, these peak fMRI responses were obtained by using the fixation trials at the beginning and at the end of the run as baselines to calculate the percentage signal changes compared with fixation (consequently, response amplitudes were higher here than what have been reported previously). The response patterns were identical between the behavioral and the fMRI measures, with both showing a response increase with increasing display set size and a response plateau at about set size 4. Specifically, responses were higher for set size 3 than 2, higher for 4 than for 3, and higher for 6 than for 3 (all values of $p < 0.05$), and responses did not different between set sizes 4 and 6 (values of $p > 0.53$). These results should not be surprising given how the superior IPS ROI was localized. Nonetheless, these results confirm and replicate previous findings and show that the superior IPS localizer was indeed successful in localizing the superior IPS ROI reported in previous studies (Todd and Marois, 2004; Xu and Chun, 2006; Xu, 2007).

Results
Researchers have shown that our ability to encode and store visual information in a temporary memory buffer called VSTM is rather limited (Phillips, 1974; Pashler, 1988; Irwin, 1992; Simons, 1996; Luck and Vogel, 1997; Rensink et al., 1997; Henderson and Hollingworth, 1999; Wheeler and Treisman, 2002; Xu, 2002a,b). VSTM integrates visual information across saccades and other interruptions as well as temporarily holds incoming or retrieved visual information to facilitate tasks such as recognition and attention-biased visual processing. As such, VSTM is an integral part of visual perception (for more supporting arguments, see Discussion). In recent fMRI and ERP studies, neural correlates of VSTM capacity limitation have been identified (Todd and

Marois, 2004; Vogel and Machizawa, 2004; Xu and Chun, 2006, 2007). Specifically, fMRI response of the superior IPS has been shown to increase with increasing display set size and plateau at the set size corresponding to the amount of object color or shape information successfully retained in VSTM (Todd and Marois, 2004; Xu and Chun, 2006, 2007; Xu, 2007). A similar result for object shapes has been observed in the LOC, a brain area involved in shape representation (Grill-Spector et al., 1999; Kourtzi and Kanwisher, 2000). fMRI response in the LOC tracks the number of unique object shapes successfully encoded and retained in VSTM (Xu and Chun, 2006, 2007, 2009; Xu, 2007, 2009).

To understand how task-irrelevant object features are processed, the present study took advantage of the fMRI response signatures in the superior IPS and the LOC in VSTM tasks. Specifically, observers were asked to encode and retain various number of object colors in VSTM and object shapes were task irrelevant. If the shape of an object is always processed together with the color of that object because of object-based selection and processing, then as the number of successfully encoded colors increases resulting in increased superior IPS response, LOC response should also increase. However, if task-irrelevant information processing is modulated by the encoding demand of the primary task, then task-irrelevant shape processing should be present when the color-encoding demand is low and attenuated when such demand is high. In a third possibility, if observers can control which object features are processed via feature-based attention (Maunsell and Treue, 2006) and task-irrelevant shapes are not processed at all, then LOC response should remain flat, independent of the color-encoding demand.

Theoretically, one can perform the reverse manipulation and ask observers to encode object shapes and examine the processing of the task-irrelevant object colors. However, set size-dependent fMRI neural response corresponding to the number of colors successfully encoded in VSTM has not been firmly established in the visual cortex. Consequently, this manipulation cannot be performed presently.

Experiment 1: VSTM color task with a short-delay period
In this fMRI study, observers were presented with a sample display containing one, two, or six objects for 200 ms. Each object had a unique color and a unique shape. After a brief blank delay of 1000 ms, they judged whether the centrally presented test probe matched in color with one of the objects in the sample display (Fig. 2 A). The encoding of object colors, but not their shapes, was task relevant.

Behavior results
To evaluate observers' behavioral performance, Cowan's K formula (Cowan, 2001) was used to transform behavioral change detection accuracies into VSTM capacity estimates (values of K) as a function of display set size. The number of object colors retained in VSTM increased with increasing display set sizes (values of $p < 0.05$ for all pairwise comparisons). Observers were able to retain maximally about three colors in VSTM (Fig. 2 B, left).

fMRI results
For each observer, fMRI responses from the VSTM task were overlaid onto that observer's independently localized superior IPS and LOC ROIs (Xu and Chun, 2006). Peak fMRI responses from these ROIs were extracted in each observer and then averaged across observers (Fig. 2 B, right).

The overall superior IPS response pattern mirrored that of the behavioral VSTM capacity measure, such that the response was
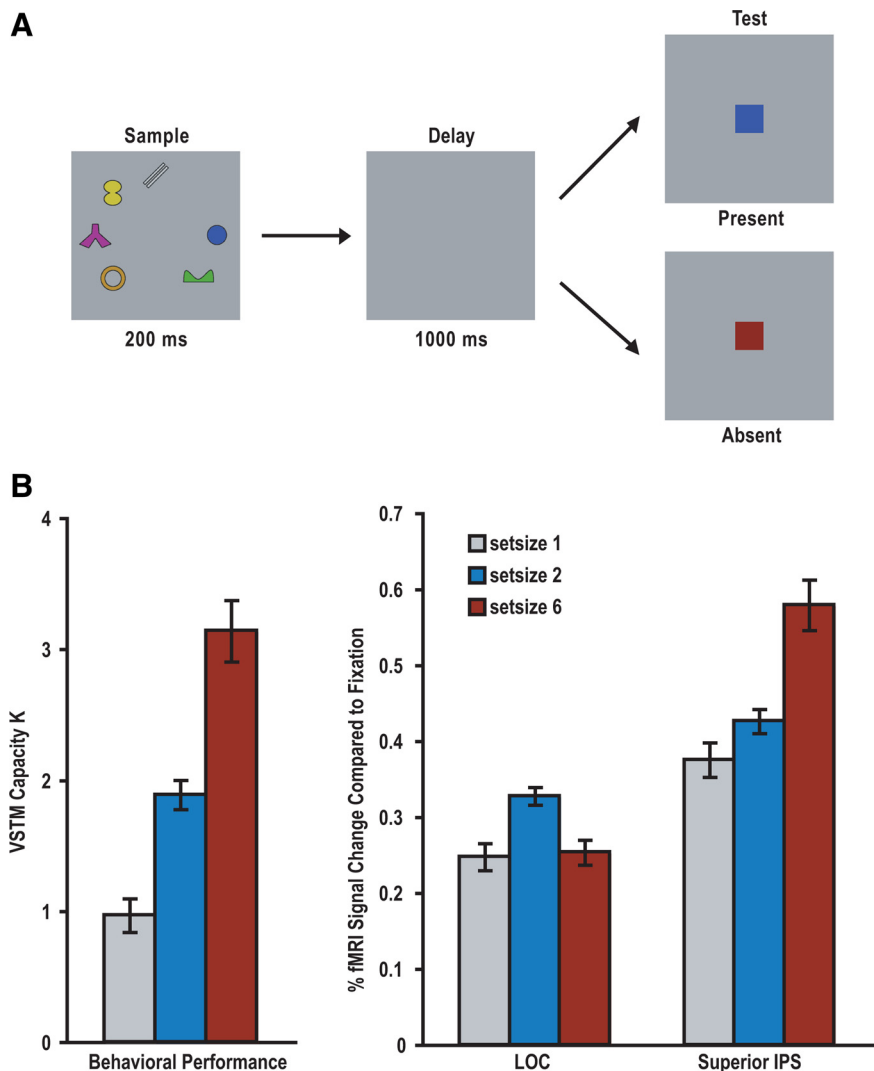
A

B

Figure 2. ***A***, An example trial from experiment 1. Observers viewed one, two, or six colored shapes in a sample display, and after a brief delay, judged whether the probe color at the center of the test display was present in the sample display. The encoding of object colors, but not their shapes, was task relevant. ***B***, Left, Behavioral performance as indicated by Cowan's *K* measures (Cowan, 2001). Color encoding increased with increasing display set sizes and observers were able to retain maximally about three colors in VSTM. Right, fMRI peak responses from the superior IPS and the LOC ROIs. The superior IPS response was low for set size 1, medium for set size 2, and high for set size 6 displays. The LOC responses for set size 1 and 6 displays did not differ from each other, and both were lower than that for set size 2 displays. Errors shown are within-subject SEs.

low for set size 1, medium for set size 2, and high for set size 6 displays (values of $p < 0.05$ for all pairwise comparisons). Because superior IPS response has been shown to track the amount of object information held in VSTM rather than overall task difficulty or the total amount of visual stimuli in a display (Todd and Marois, 2004; Xu and Chun, 2006; Xu, 2007), this result indicated that more color information was encoded and retained in the superior IPS when display set size increased from one to two to six.

The overall LOC response pattern differed from that of the superior IPS and the behavioral measure: LOC responses for set size 1 and 6 displays did not differ from each other ($p > 0.86$), and were both lower than that for set size 2 displays (values of $p < 0.05$). Thus, more shape information was processed when the color-encoding demand increased from one to two, showing an object-based processing benefit for a task-irrelevant feature and replicating previous findings (O'Craven et al., 1999). As the color-encoding demand increased from two to six, however, shape processing seemed to be attenuated and possibly sup-

pressed, with the presence of six object shapes elicited less LOC response than that of two such shapes. This is in sharp contrast to the increase in response seen in the superior IPS ($p < 0.001$ for the interaction between the two brain regions and the two set sizes). Note that, when these object shapes were task relevant in another study (Xu and Chun, 2006), LOC response was capable of increasing from set size 2 to 6. The absence of such response increase here suggests that object-based shape processing was modulated by the encoding demand of the primary task, such that not all the irrelevant shapes can be processed together with their colors when the color-encoding demand was high.

Experiment 2: VSTM color task with a long-delay period

A VSTM task contains three distinctive stages: encoding, maintenance, and retrieval/comparison. Where does the object-based processing occur? Is it present during both feature encoding and maintenance, or only during encoding? These questions could not be answered in experiment 1 because of the brief delay period used and the lag in fMRI responses. To address this question, this experiment repeated the design of experiment 1 but with a VSTM delay period of 8300 ms instead of 1000 ms following Xu and Chun (2006), experiment 3 and Todd and Marois (2004). When object shapes were directly attended in a previous study, persistent set size-dependent delay period activity was observed in the LOC (Xu and Chun, 2006). Thus, whether such persistent activity is present or not in the LOC would indicate whether the processing of task-irrelevant shape information is long-lasting or short-lived.

Behavior results

As in experiment 1, behavioral performance was transformed into VSTM capacity estimates. The number of object colors retained in VSTM increased with increasing display set sizes (values of $p < 0.05$ for all pairwise comparisons). Observers were able to retain maximally about three colors in VSTM (Fig. 3, left), similar to what was found in experiment 1.

fMRI results

As in experiment 1, for each observer, fMRI responses from the VSTM task were overlaid onto that observer's independently localized superior IPS and LOC ROIs, and the resulting time courses were averaged and plotted in Figure 3, right. There were two peaks in the fMRI time course plots, corresponding to the presentation of the sample and the test displays, respectively (Fig. 3, right). Each time point from the rise of the first peak to the rise of the second peak (i.e., from time points 10.5 to 19.5 s) was examined, and the resulting pairwise comparisons for the three display conditions were reported in Tables 1 and 2.

In the superior IPS, significant pairwise differences were observed around the first peak as well as between the two peaks in the time course data. This indicates that different amount of color information was represented in the superior IPS from the three display conditions during both VSTM encoding and maintenance (Fig. 3, right; Table 1), consistent with behavioral VSTM capacity measures as well as previous fMRI findings (Todd and Marois, 2004; Xu and Chun, 2006).

In the LOC, significant pairwise results were only observed around the first peak of the time course data (Fig. 3, right; Table 2), showing that differences between the display conditions were only present during VSTM encoding but not during the delay period. Specifically, during encoding, responses were lower for set size 1 than for either set size 2 or 6 displays, with no difference between the latter two. That is, more shape information was encoded when the color-encoding demand increased from one to two, but this effect plateaued when the color-encoding demand increased further from two to six. Thus, the object-based processing benefit is modulated by the encoding demand of the primary task, such that task-irrelevant shapes could not always be encoded together with their colors, consistent with the findings of experiment 1.

During the delay period, none of the pairwise comparisons reached significance in the LOC, indicating that the object-based processing effect was transient and not retained over a long delay period in this brain area. Note that this is not because LOC response could not be sustained during the VSTM delay period. When shapes were directly attended, LOC response has been shown to be sustained during the delay period (Xu and Chun, 2006). Moreover, although the encoding signal was equivalent for the set size 1 displays in both the superior IPS and the LOC, the former remained above baseline during the delay period, whereas the latter returned to baseline in the same time period. Thus, the differences in results between the two brain regions cannot be attributable to a greater signal during encoding returning to baseline more slowly (in the superior IPS) than a smaller signal (in the LOC). Instead, the response pattern indicated that there was indeed sustained delay period activity in the superior IPS, but not in the LOC.

In experiment 1, LOC response was lower for the set size 6 than for the set size 2 displays. This difference, however, was not found in this study. In experiment 1, because of the brief delay period used, observers had to quickly consolidate the encoded color information into VSTM and be ready to perform change detection on the upcoming color probe. The urgency of these closely spaced operations might have caused less shape processing or even suppression in the LOC to decrease the competition between color and shape encoding at set size 6 when all the processing resources were needed for color encoding. In this study, there was ample time to consolidate the encoded color information into VSTM. Moreover, even if the irrelevant shape features were encoded, they would quickly decay if no attempt was made to actively maintain them in VSTM. As such, it might not be necessary to actively suppress shape en-
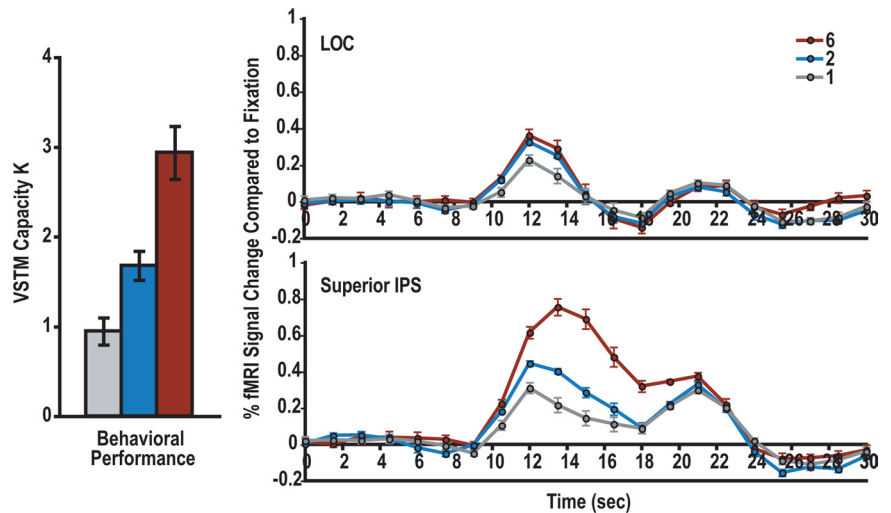

Figure 3. Results from experiment 2. Left, Behavioral performance as indicated by Cowan's K measures (Cowan, 2001). As in experiment 1, color encoding increased with increasing display set sizes and observers were able to retain maximally about three colors in VSTM. Right, fMRI response time courses from the superior IPS and the LOC ROIs. The two peaks in the time course plots correspond to the presentation of the sample and the test displays, respectively. In the superior IPS, significant pairwise differences were observed around the first peak as well as between the two peaks in the time course data. In the LOC, significant pairwise comparisons were only observed around the first peak of the time course data, with the LOC response being lower for set size 1 than for either set size 2 or 6 displays and with no difference between the latter two display conditions. Errors shown are within-subject SEs.

Table 1. Values of *p* for pairwise comparisons in experiment 2 for superior IPS responses

Time points (s)	1 versus 2	2 versus 6	1 versus 6
10.5	0.050*	0.036*	0.034*
12.0	0.005**	0.005**	0.002**
13.5	0.002**	0.000**	0.000**
15.0	0.014*	0.001**	0.001**
16.5	0.145	0.008**	0.004**
18.0	0.942	0.004**	0.003**
19.5	0.748	0.000**	0.001**

*$p < 0.05$.
**$p < 0.01$.

Table 2. Values of *p* for pairwise comparisons in experiment 2 for LOC responses

Time points (s)	1 versus 2	2 versus 6	1 versus 6
10.5	0.063 †	0.784	0.079 †
12.0	0.046*	0.219	0.006**
13.5	0.086 †	0.157	0.030*
15.0	0.949	0.768	0.843
16.5	0.362	0.769	0.342
18.0	0.143	0.600	0.240
19.5	0.588	0.528	0.332

*$p < 0.05$.
**$p < 0.01$.
†$0.05 < p < 0.10$.

coding in the LOC. This may explain the discrepancies seen in these two studies.

Experiment 3: VSTM short-delay task with color and shape repetition

There were two possible concerns for the results of experiments 1 and 2. The first concern is that, because it was more attentional demanding to encode the colors of two than one shape, this attentional elevation could have increased responses in multiple brain areas without necessarily signaling the encoding of more shapes in the LOC. In other words, the increase in LOC response
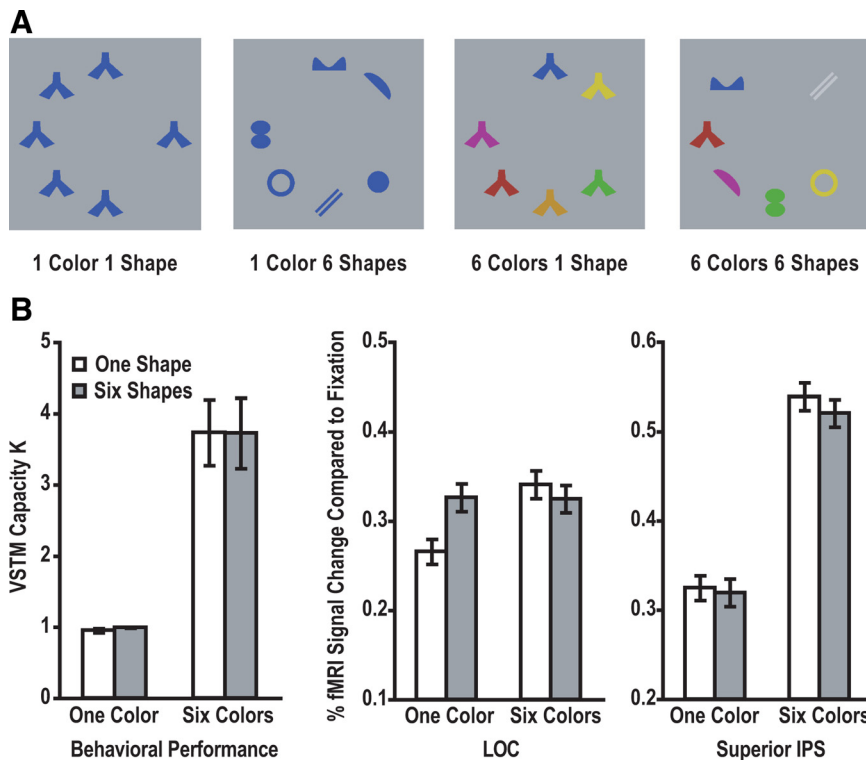
Figure 4. *A*, Sample stimuli used in experiment 3. Each sample display consisted of six colored shapes, which contained either a total of one unique shape and one unique color, a total of one unique shape and six unique colors, a total of six unique shapes and one unique color, or a total of six unique shapes and six unique colors. *B*, Left, Behavioral performance as indicated by Cowan's *K* measures (Cowan, 2001). The number of object colors retained in VSTM increased as the number of colors in the display increased from one to six. There was no interaction between the color-encoding load and shape identity. *B*, Middle and Right, fMRI peak responses from the superior IPS and the LOC ROIs. Superior IPS response mirrored the behavioral VSTM capacity measure, showing no interaction between color-encoding load and shape identity. In the LOC, when only one color was encoded, response was lower for six identical than for six different shapes; and when six colors were encoded, however, there was no effect of shape identity. There was a significant interaction between the color-encoding load and shape identity. Errors shown are within-subject SEs.

sentation. When attention is directed to one of the multiple stimuli via either top-down or bottom-up process, however, neural competition between that and the other stimuli can be successfully overcome (Desimone and Duncan, 1995; Kastner et al., 1998; Reynolds et al., 1999; Beck and Kastner, 2005). Furthermore, when attention is directed to multiple items simultaneously, I found in a previous fMRI study that four different shapes invoked a much stronger LOC response than four identical shapes (Xu, 2009), opposite of the neural competition response reported by Beck and Kastner (2007). Thus, whereas multiple unattended objects may evoke neural competition and result in poorer representation, multiple attended objects can evoke a stronger neural response and result in corepresentation.

In experiments 1 and 2, because attention was directed to each colored shape to encode its color and thus each object was attended, neural competition should have been overcome and resulted in a stronger LOC response for six than for two shapes, yielding a result similar to that of Xu and Chun (2006). The failure of obtaining such a result indicated that it seemed unlikely that neural competition between the items played a role in these two studies.

To examine the attentional account and to confirm that neural competition did not play a role in experiments 1 and

when set size increased from 1 to 2 might entirely be attributable to this attentional elevation rather than object-based encoding of the task-irrelevant shapes at low color-encoding load. This account, however, cannot explain why LOC response failed to increase further when the attentional demand for color encoding increased continuously from set size 2 to 6.

The second concern for the results of experiments 1 and 2 is that, because more objects were present at set size 6 than at set size 2, there might have been a stronger neural competition for stimulus representation at set size 6, and this could reduce the LOC response. In other words, the failure of LOC response to increase further from set size 2 to 6 in both experiments might have been attributable to neural competition at a higher set size, rather than a reduced object-based encoding at the higher encoding load for the task-relevant object feature.

Both neurophysiological and fMRI studies have shown that multiple objects compete for neural representation (Desimone and Duncan, 1995; Kastner et al., 1998; Reynolds et al., 1999; Beck and Kastner, 2005, 2007, 2009). For example, Kastner et al. (1998) showed that four unattended simultaneously presented peripheral objects activated the human early visual areas (V1–V4 and TEO) significantly weaker than the same objects presented sequentially. Using a similar manipulation, Beck and Kastner (2007) found that, in the human area V4, four unattended peripheral objects that were all different had a lower fMRI response than four unattended peripheral objects there were all identical, showing that perceptual grouping can overcome neural competition for repre-

2, in this experiment, observers attended the colors of six shapes. These shapes contained either a total of one unique shape and one unique color, a total of one unique shape and six unique colors, a total of six unique shapes and one unique color, or a total of six unique shapes and six unique colors (for some examples, see Fig. 4*A*).

If task-irrelevant object features are not encoded at low color-encoding load, LOC response would be identical when one unique color was encoded regardless of whether one or six unique shapes were present. However, if such object-based encoding does exist, then one color appearing on six unique shapes would evoke a stronger LOC response than one color appearing on six identical shapes. At high color-encoding load, both accounts would predict that the LOC response would be similar regardless of whether one or six unique shapes were present.

If neural competition plays a role in representing multiple shapes, according to Beck and Kastner (2007), six different shapes would not invoke a stronger response than six identical shapes in the LOC, independent of the color-encoding load. However, if object-based encoding of the task-irrelevant feature is modulated by the encoding demand of the task-relevant color, then six different shapes would invoke a stronger LOC response than six identical shapes only at low, but not at high, color-encoding load.

Behavior results
As in experiments 1 and 2, behavioral performance was transformed into VSTM capacity estimates (Fig. 4, bottom left). The

number of object colors retained in VSTM increased as the number of colors in the display increased from one to six (value of $p < 0.001$ for all pairwise comparisons between one-color and six-color displays). When observers encoded one color from the display, performance did not differ between whether six identical or six different shapes were present ($F < 1$); similarly, when observers encoded six colors from the display, performance did not differ either between whether six identical or six different shapes were present ($F < 1$). There was no interaction between the color-encoding load and shape identity ($F < 1$).

fMRI results
As in experiment 1, for each observer, fMRI responses from the VSTM task were overlaid onto that observer's independently localized superior IPS and LOC ROIs. Peak fMRI responses from these ROIs were extracted in each observer and then averaged across observers (Fig. 4, bottom right).

The overall superior IPS response pattern mirrored that of the behavioral VSTM capacity measure. Response was lower when only one than when six colors were present in the displays (values of $p < 0.05$ for all pairwise comparisons between one-color and six-color displays), indicating that more color information was encoded and retained in the superior IPS when the color-encoding load increased from one to six. Whether six identical or six different shapes were present in the display, however, had no effect on color encoding ($F < 1$), and there was no interaction between the color-encoding load and shape identity ($F < 1$). This shows that the task-irrelevant shape information was not retained in the superior IPS regardless of the color-encoding load.

The LOC response pattern differed from that of the superior IPS. Specifically, there was a significant interaction between the color-encoding load and shape identity ($p < 0.05$). When only one color was present in the display, response was lower for six identical than for six different shapes ($p < 0.05$). Note that this cannot be because fewer shapes were attended in the identical than in the all different condition. To do so, one would have already attended and encoded all the shapes to know whether they were all identical or not. When six colors were present in the display, however, response did not differ between whether six identical or six different shapes were present ($F < 1$). Note that this is not because LOC response could not increase any higher. When a blocked fMRI design was used and when shapes were directly attended, LOC response amplitude could easily double (Xu, 2008) compared with what was reported here with an event-related design. These results show that the task-irrelevant shapes were encoded by the LOC at the low, but not at the high, color-encoding load, consistent with the findings from experiments 1 and 2.

Given that the same number of shapes was present in all the conditions, the LOC result further indicates that neural competition between items did not play a role in the present study. Interestingly, when only one unique shape was present, LOC response was higher for encoding six than for encoding just one color ($p < 0.05$). Because it was more attentional demanding to encode more colors, this attentional elevation seemed to have modulated responses in the LOC. Thus, the LOC response from experiments 1 and 2 for set size 2 displays could have contained both attention and shape encoding-related activations. Even though the attentional demand for color encoding increased further from set size 2 to 6, LOC response failed to increase further in those experiments. This suggests that shape encoding must have been greatly suppressed at set size 6 to cancel entirely (experiment

1) or partially (experiment 2) any attention-induced LOC response increase.

To summarize, in the present experiment, when the attentional demand was equated at low color-encoding load, six different shapes evoked a stronger LOC response than six identical shapes. At high color-encoding load, although LOC response was capable of rising much higher, six different shapes did not evoke a stronger LOC response than six identical shapes. Thus, whereas the object-based encoding of the task-irrelevant shapes was present at low color-encoding load, it was absent at high color-encoding load, replicating the results of experiments 1 and 2.

Discussion
To determine when object-based processing may or may not occur, the present fMRI study varied the overall task encoding load and examined the processing of object shapes when attention was directed to object colors. Specifically, the processing of unattended object shapes was assessed by measuring fMRI responses in the LOC, a brain area involved in object shape representation. In three experiments, it was found that, whereas object-based processing was present at low color-encoding load, it was attenuated or even suppressed at high color-encoding load. This effect could not be explained by an attentional effect or stronger neural competition at higher display set size (see experiment 3). Instead, these results show that object-based processing is not always free and automatic. But rather, it depends on the encoding demand of the task-relevant object feature.

Interestingly, even when the task-irrelevant shapes were encoded by the LOC, experiment 3 showed that this information was not processed and retained in the superior IPS, a brain area involved in VSTM information encoding and storage (Todd and Marois, 2004; Xu and Chun, 2006; Xu, 2007). Thus, although the object-based shape encoding may be automatic and obligatory under low color-encoding load in higher visual areas, this is not the case for information processing in the parietal cortex. The superior IPS seems to process only the task-relevant features, consistent with a previous fMRI result [Xu and Chun (2006), experiment 1] and related neurophysiological findings (Toth and Assad, 2002; Freedman and Assad, 2006).

Another major finding of the present study is that the object-based processing of the task-irrelevant feature is transient. Experiment 2 showed that the processing of the task-irrelevant shape features was short lasting and quickly decayed when no attempt was made to sustain it. Thus, although object-based shape processing at low color-encoding load may be initially automatic and obligatory, observers can exert control over which object information is retained for a prolonged period of time.

Object-based encoding: under what condition does it occur?
Although some previous studies have reported object-based processing, others have failed to find such benefit. How do we resolve these discrepancies? O'Craven et al. (1999) asked observers to perform a simple one-back repetition detection on the task-relevant object feature. Because the primary task was relatively easy and the stimulus encoding load was low, consistent with the present results, O'Craven et al. found object-based processing for the task-irrelevant features.

Compared with O'Craven et al. (1999), both Woodman and Vogel (2008) and Serences et al. (2009) used much harder tasks. In the behavioral study of Woodman and Vogel (2008), initial stimulus presentation was always masked and observers could retain no more than two of the three objects shown. In their ERP study, performance again was not at ceiling. Woodman and

Vogel (2008) thus might have used a task with a high stimulus encoding load that completely suppressed the encoding of the task-irrelevant feature right from the onset of the stimulus display. This is consistent with the LOC results from experiment 1 for set size 6 displays. Serences et al. (2009) titrated task difficulty for each observer until a criterion level of performance of ~75% correct was reached. Moreover, this study averaged neural responses over a long delay period. Thus, under a higher encoding load for the task-relevant stimulus and over a long delay period, the object-based processing benefit was not found. This is again consistent with the present findings.

Unlike the studies by Woodman and Vogel (2008) and Serences et al. (2009), however, the present study finds that, when the encoding demand of the task-relevant feature was low, automatic and obligatory object-based processing of the task-irrelevant feature did exist during the initial stimulus encoding. Thus, compared with previous findings, the present study delineates conditions in which object-based processing may or may not occur and provides a plausible account for the discrepancies among previous studies.

Presently, it is unknown whether task-relevant and task-irrelevant features are encoded with the same resolution during object-based processing. When it is task relevant, object shapes need to be encoded in sufficient resolution to support later memory recognition; when it is task irrelevant, however, it is unclear whether shapes are still encoded in such fine resolution, or whether they are only coarsely encoded. If the latter is true and task-irrelevant coarse shape features are encoded together with the task-relevant color features, then a faster ERP time course (reflecting feature encoding and consolidation) could be obtained for encoding color than for encoding shape. This may contribute to the ERP results reported by Woodman and Vogel (2008). Additional studies are needed to address this question to allow us to fully understand how task-irrelevant object features are encoded during object-based processing.

Object-based encoding and perceptual load manipulation

Experiment 2 showed that the processing of the task-irrelevant shape features was short lasting and quickly decayed when no attempt was made to sustain it. This suggests that the object-based encoding in the LOC reflects perceptual and attentional processes, which appear to be automatic, rather than memory processes, which appear to be under voluntary control. Because this object-based encoding of the task-irrelevant features in the LOC depends on the encoding load of the task-relevant feature, it is similar to the effect described by the perceptual load theory, which reports that increasing the perceptual load of target perception decreases distractor processing (Lavie and Tsal, 1994; Lavie, 1995, 2005; de Fockert et al., 2001; Handy et al., 2001; Lavie et al., 2004).

The present study, however, differs from previous studies on perceptual load in an important way. In previous studies, there was usually a spatial separation between target and distractors. It has been argued that "reduced distractor processing in high perceptual load might be attributable to narrowing of a spatial attention window around the target space, effectively excluding stimuli outside it" (Lavie, 2005). Here, the task-irrelevant shape features were on the same objects containing the task-relevant color features. If anything, task-irrelevant shapes should have been attended more at the high than at the low color-encoding load. And yet, shape encoding was attenuated at high load. This finding cannot be explained by size differences of the visual attentional window at different loads. Rather, the capacity limita-

tion of the central processing resources may ultimately determine whether task-irrelevant object features are encoded or not.

The distinction between perception and VSTM

VSTM is an integral part of visual perception (Xu, 2002b). This is because VSTM integrates visual information across saccades and other interruptions as well as temporarily holds incoming or retrieved visual information to facilitate tasks such as recognition and attention biased visual processing. As such, it is impossible to define where perception ends and where the encoding of VSTM starts. Even without exerting control or effort, we often retain and remember for a short duration part of what we have just perceived. In the cognitive psychology literature, paradigms used to study perception and attention often involve recall or recognition, which necessarily engages VSTM. For example, to study object-based attention, Duncan (1984) asked observers to report a briefly presented and masked visual stimulus. It is thus impossible to conclude whether results obtained from such paradigm were unique to perception and attention. In the cognitive neuroscience literature, studies have shown that VSTM information maintenance recruits the same brain areas involved in perception (Jonides et al., 2005; Pasternak and Greenlee, 2005). In this regard, VSTM can be viewed as an extension of perception (i.e., extending the representation formed during perception for a longer period). The present results and conclusions thus speak to object-based encoding in both perception and VSTM.

Although it is difficult to define the boundary between perception and VSTM, these two cognitive operations do differ in an important way. The present study shows that, although the encoding of the task-irrelevant object feature may be automatic at the low encoding load of the task-relevant object feature, the maintenance of this information in VSTM for a prolonged period, however, is under voluntary control. Without such control and effort, the processing of the task-irrelevant object feature was short lasting and quickly decayed.

In summary, the present study found that object-based encoding of the task-irrelevant features does occur, as reported previously; but it is transient and its magnitude is determined by the overall encoding demand of the task-relevant object feature.

References

Allport DA (1971) Parallel encoding within and between elementary stimulus dimensions. Percept Psychophys 10:104–108.

Beck DM, Kastner S (2005) Stimulus context modulates competition in human extrastriate cortex. Nat Neurosci 8:1110–1116.

Beck DM, Kastner S (2007) Stimulus similarity modulates competitive interactions in human visual cortex. J Vis 7:19.1–12.

Beck DM, Kastner S (2009) Top-down and bottom-up mechanisms in biasing competition in the human brain. Vision Res 49:1154–1165.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

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.

de Fockert JW, Rees G, Frith CD, Lavie N (2001) The role of working memory in visual selective attention. Science 291:1803–1806.

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

Duncan J (1984) Selective attention and the organization of visual information. J Exp Psychol Gen 113:501–517.

Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. Nature 443:85–88.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron 24:187–203.

Handy TC, Soltani M, Mangun GR (2001) Perceptual load and visuocortical processing: event-related potentials reveal sensory-level selection. Psychol Sci 12:213–218.

Henderson JM, Hollingworth A (1999) High-level scene perception. Annu Rev Psychol 50:243–271.

Irwin DE (1992) Memory for position and identity across eye movements. J Exp Psychol Learn 18:307–317.

Jonides J, Lacey SC, Nee DE (2005) Processes of working memory in mind and brain. Curr Dir Psychol Sci 14:2–5.

Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. Science 282:108–111.

Kourtzi Z, Kanwisher N (2000) Cortical regions involved in perceiving object shape. J Neurosci 20:3310–3318.

Lavie N (1995) Perceptual load as a necessary condition for selective attention. J Exp Psychol Hum Percept Perform 21:451–468.

Lavie N (2005) Distracted and confused?: selective attention under load. Trends Cogn Sci 9:75–82.

Lavie N, Tsal Y (1994) Perceptual load as a major determinant of the locus of selection in visual attention. Percept Psychophys 56:183–197.

Lavie N, Hirst A, de Fockert JW, Viding E (2004) Load theory of selective attention and cognitive control. J Exp Psychol Gen 133:339–354.

Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. Nature 390:279–281.

Maunsell JH, Treue S (2006) Feature-based attention in visual cortex. Trends Neurosci 29:317–322.

O'Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. Nature 401:584–587.

Pashler H (1988) Familiarity and visual change detection. Percept Psychophys 44:369–378.

Pasternak T, Greenlee MW (2005) Working memory in primate sensory systems. Nat Rev Neurosci 6:97–107.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442.

Phillips WA (1974) On the distinction between sensory storage and short-term visual memory. Percept Psychophys 16:283–290.

Rensink RA, O'Regan JK, Clark JJ (1997) To see or not to see: the need for attention to perceive changes in scenes. Psychol Sci 8:368–373.

Reynolds JH, Chelazzi L, Desimone R (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci 19:1736–1753.

Scholl BJ (2001) Objects and attention: the state of the art. Cognition 80:1–46.

Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. Psychol Sci 20:207–214.

Simons DJ (1996) When object representations fail. Psychol Sci 7:301–305.

Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain (Rayport M, translator). New York: Thieme Medical.

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

Toth LJ, Assad JA (2002) Dynamic coding of behaviourally relevant stimuli in parietal cortex. Nature 415:165–168.

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 (2001) Storage of features, conjunctions, and objects in visual working memory. J Exp Psychol Hum Percept Perform 27:92–114.

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.

Wheeler ME, Treisman AM (2002) Binding in short-term visual memory. J Exp Psychol Gen 131:48–64.

Woodman GF, Vogel EK (2008) Selective storage and maintenance of an object's features in visual working memory. Psychon Bull Rev 15:223–229.

Xu Y (2002a) Limitations of object-based feature encoding in visual short-term memory. J Exp Psychol Hum Percept Perform 28:458–468.

Xu Y (2002b) Encoding color and shape from different parts of an object in visual short-term memory. Percept Psychophys 64:1260–1280.

Xu Y (2007) The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. J Neurosci 27:11676–11686.

Xu Y (2008) Representing connected and disconnected shapes in human inferior intra-parietal sulcus. Neuroimage 40:1849–1856.

Xu Y (2009) Distinctive neural mechanisms supporting visual object individuation and identification. J Cogn Neurosci 21:511–518.

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

Xu Y, Chun MM (2007) Visual grouping in human parietal cortex. Proc Natl Acad Sci U S A 104:18766–18771.

Xu Y, Chun MM (2009) Selecting and perceiving multiple visual objects. Trends Cogn Sci 13:167–174.