

Neural Representation of Targets and Distractors during Object Individuation and Identification

Su Keun Jeong and Yaoda Xu

Abstract

■ In many everyday activities, we need to attend and encode multiple target objects among distractor objects. For example, when driving a car on a busy street, we need to simultaneously attend objects such as traffic signs, pedestrians, and other cars, while ignoring colorful and flashing objects in display windows. To explain how multiple visual objects are selected and encoded in visual STM and in perception in general, the neural object file theory argues that, whereas object selection and individuation is supported by inferior intraparietal sulcus (IPS), the encoding of detailed object features that enables object identification is mediated by superior IPS and higher visual areas such as the lateral occipital complex (LOC). Nevertheless, because task-irrelevant distractor objects were never present in previous studies, it is unclear how distractor objects would impact neural responses related to target object individuation and identification. To address this question, in two fMRI experiments, we

asked participants to encode target object shapes among distractor object shapes, with targets and distractors shown in different spatial locations and in different colors. We found that distractor-related neural processing only occurred at low, but not at high, target encoding load and impacted both target individuation in inferior IPS and target identification in superior IPS and LOC. However, such distractor-related neural processing was short-lived, as it was only present during the visual STM encoding but not the delay period. Moreover, with spatial cuing of target locations in advance, distractor processing was attenuated during target encoding in superior IPS. These results are consistent with the load theory of visual information processing. They also show that, whereas inferior IPS and LOC were automatically engaged in distractor processing under low task load, with the help of precuing, superior IPS was able to only encode the task-relevant visual information. ■

INTRODUCTION

Encoding, retaining, and retrieving visual information relevant to behavior and thoughts are some of the most fundamental human cognitive abilities. Over the past six decades, pioneered by human neuropsychological studies on patients such as H.M. (Corkin, 1968; Milner, Corkin, & Teuber, 1968; Scoville & Milner, 1957; see also Corkin, 2002; Corkin, Amaral, González, Johnson, & Hyman, 1997), many insights have been gained regarding the role of the medial-temporal lobe in mediating information retention in long-term memory. Meanwhile, how visual information is first perceived and retained in visual STM (VSTM) has been linked to the functions of the pFC and the parietal cortex (e.g., Xu & Chun, 2006; Todd & Marois, 2004; Ungerleider, Courtney, & Haxby, 1998; Goldman-Rakic, 1995).

In one study using fMRI, Xu and Chun (2006) asked participants to encode multiple object shapes into VSTM. They found that responses in inferior intraparietal sulcus (IPS) increased with increasing object number and plateaued at about set size 4 regardless of object complexity. In addition, they found that responses from superior IPS

and lateral occipital complex (LOC, an object shape area; see Malach et al., 1995) increased with set size and plateaued at about the maximal number of objects held in VSTM (equal or less than four) as determined by object complexity. On the basis of these and other related findings, Xu and Chun proposed the neural object file theory and argued that, in VSTM as well as in perception in general, object individuation is supported by inferior IPS and object identification is mediated by superior IPS and higher object processing regions such as LOC (see also Xu, 2007, 2008, 2009; Xu & Chun, 2007, 2009). Here, object individuation refers to the selection of objects via their spatial locations, whereas object identification refers to the encoding of detailed object featural information. These neural findings are in line with previous behavioral findings and theories regarding how the visual system selects and encodes multiple objects through individuation and identification processes (Pylyshyn, 1989, 1994; Kahneman, Treisman, & Gibbs, 1992).

Nevertheless, because only targets were included in previous studies (Xu, 2007, 2009; Xu & Chun, 2006, 2009), it is unclear how the neural mechanisms mediating object individuation and identification would operate in the presence of task-irrelevant distractors. Understanding the impact of distractors during object individuation and

Harvard University

identification is essential if we want to generalize laboratory findings to real-world object perception, as irrelevant visual information is always present in everyday vision.

How distractors are filtered out by our visual system has been examined by research dated back to the 1950s. The early selection view argues that the visual system can select targets and ignore distractors very early on during visual processing (Moray, 1959; Broadbent, 1958). According to this view, the presence of distractors should have minimum impact on the neural responses mediating visual object individuation and identification. Alternatively, the late selection view argues that our visual system can individuate or even identify distractors (Luck, Vogel, & Shapiro, 1996; Deutsch & Deutsch, 1963). According to this view, the presence of distractors would significantly impact neural substrates supporting object individuation and identification. A third possibility is that the processing of distractors depends on the available resources. Accordingly, irrelevant information is processed only when the main task is relatively easy and does not consume all the available resources (Lavie, 2005; Lavie & Tsai, 1994; see also Yi, Woodman, Widders, Marois, & Chun, 2004). This view would predict that distractors will only be processed and impact neural responses when the demand for object individuation and identification is low. Indeed, when Xu (2010) examined the encoding of two features from the same object, with one being task-relevant and the other task-irrelevant, she found that object-based encoding of task-irrelevant object features only occurred when the demand to encode task-relevant object features was low. Moreover, such object-based processing was short-lived and was not sustained over a long delay period.

In this study, we investigated the impact of task-irrelevant distractors on the neural mechanisms supporting object individuation and identification when targets and distractors appeared in different spatial locations. In Experiment 1, we asked participants to encode target shapes among distractor shapes in a VSTM task. A long delay period was used to allow us to separately examine encoding-, delay-, and retrieval-associated neural responses. In Experiment 2, we asked whether top-down attention could modulate distractor processing during object individuation and identification. By using either neutral or valid location cues, we tested whether distractor processing could be excluded when participants knew target locations in advance.

EXPERIMENT 1

In this experiment, we examined the impact of distractors on object individuation and identification during both the VSTM encoding and delay periods. We varied the target load by presenting either one or four target shapes in one color and varied the distractor load by presenting either zero or three distractor shapes in a different color. We measured neural responses in independently defined inferior IPS, superior IPS, and LOC ROIs. The early selection theory would predict that distractors would be

processed regardless of the encoding demand. The late selection theory, on the other hand, would predict that distractors would be filtered out during the encoding period. Lastly, the load theory would predict that processing of distractors would depend on the target encoding load.

Methods

Participants

Twelve paid participants (seven women) were recruited from the Harvard University community (mean age = 23.83 years, $SD = 4.87$ years) with informed consent, which was approved by the institutional review board of Harvard University. All of them were right-handed and had normal or corrected-to-normal visual acuity. One additional participant was tested but was excluded from further analysis due to excessive head motion (more than 5 mm).

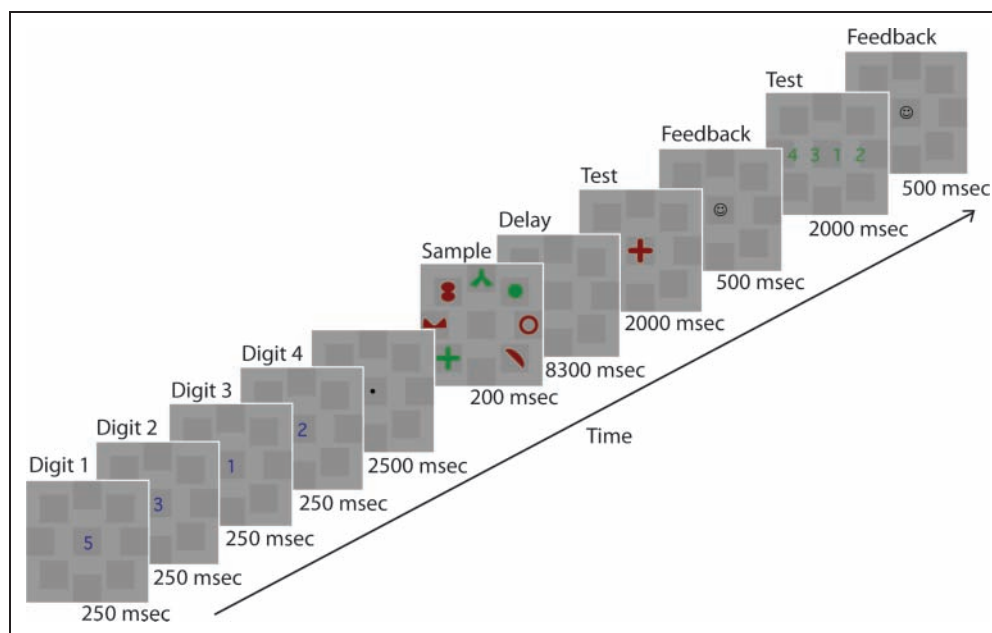
Main Experimental Design

The participants were asked to remember target shapes among distractor shapes presented briefly around the central fixation. After an extended delay, they judged whether a probed shape matched one of the remembered target shapes by pressing either the “match” or the “no-match” key (see Figure 1 for an illustration of the trial sequence). A match occurred in half of the trials. Targets and distractors were shown in different colors to facilitate target selection, with half the participants having red targets and green distractors and the other half having the reverse color assignment.

There were four conditions: 1 target with 0 distractors (1T), 1 target with 3 distractors (1T + 3D), 4 targets with 0 distractors (4T), and 4 targets with 3 distractors (4T + 3D). All stimuli appeared on a light gray background. To prevent grouping, eight dark gray squares were also presented as placeholders and marked all the possible locations for which targets and distractors could appear (see Figure 1; see also Xu, 2009). Eight different target and distractor shapes were used (see Xu & Chun, 2006), each subtended approximately $2.74^\circ \times 2.74^\circ$. The size of the entire display was $11.8^\circ \times 11.8^\circ$.

To prevent participants from verbally encoding the shapes, in addition to the VSTM shape task, they were required to remember and rehearse four digits throughout each VSTM trial by comparing whether four digits presented sequentially at the beginning of each trial matched those presented simultaneously at the end of each trial. Inferior IPS has been shown to track the number of objects presented at different spatial locations (up to four locations, see Xu, 2009; Xu & Chun, 2006). As such, given the 6-sec lag in hemodynamic response, simultaneous presentation of the four digits at different spatial locations may saturate inferior IPS response before the presentation of the target and distractor stimuli (which occurred 2.5 sec after the digit presentation). For this reason, digits

Figure 1. An example trial of Experiment 1. Participants were asked to remember target shapes (in red) among distractor shapes (in green). After an extended delay, they judged whether the probed shape matched one of the remembered target shapes by pressing the appropriate response button. A match occurred in half of the trials. Target and distractor color assignment was balanced across different participants. To discourage grouping, eight dark gray squares were also present as placeholders and marked all possible locations that targets and distractors could appear. To prevent verbal encoding of the shapes, in addition to the VSTM shape task, four digits were shown sequentially at the beginning of each trial. Participants were asked to remember and rehearse these digits and then judge whether the same four digits were shown at the end of the trial.



were presented sequentially, rather than simultaneously, at the beginning of each trial. Each trial lasted 18 sec and contained the following: a fixation period (1000 msec), a sequential presentation of four digits (250 msec each), a fixation period (2500 msec), a sample shape display (200 msec), a delay period (8300 msec), a test shape display (2000 msec), a shape response feedback (500 msec), a test digit display (2000 msec), and a digit response feedback (500 msec; Figure 1). The participants were instructed to maintain fixation during the trial. With a counterbalanced trial history design (see Xu & Chun, 2006; Todd & Marois, 2004), each run contained a total of 27 trials, including five trials for each stimulus condition, five fixation trials, and two filler trials. Fixation trials contained the digit task without the VSTM shape task (the shape task was replaced by a fixation dot). Filler trials were included to balance trial history, with one appearing at the beginning and one at the end of the run. Filler trials were excluded during data analysis. Each participant completed four or five runs, with each run lasting 8 min and 15 sec.

Localizer Design

To ensure that the ROIs we localized were involved in processing the specific visual stimuli used in the main experiment, the shapes from the main experiment appeared in the same size and eccentricity in all the ROI localizers described below as they did in the main experiment.

To localize the superior IPS region that closely tracks the amount of visual information retained in VSTM, we conducted an independent shape VSTM experiment similar to that of Xu and Chun (2006). Specifically, participants were asked to remember one, two, three, four, or six black object shapes presented briefly around the central fixation. After a short delay, a probe shape appeared at fixation and required participants to make a probe match/no-match judgment. The probe matched one of the remembered shapes in half of the trials. Each trial lasted 6 sec and contained the following: a fixation period (1000 msec), a sample display (200 msec), a delay period (1000 msec), a test shape display/response period (2500 msec), and a feedback (1300 msec). The sizes of the individual object shape and the whole display were identical to those used in the main VSTM experiment. With a counterbalanced trial history design, there were 12 stimulus trials for each set size condition as well as 12 fixation trials in which only a fixation dot appeared during the 6-sec trial period. Three filler trials were added to the beginning, and one filler trial was added to the end of each run for practice and trial history balancing purposes. These filler trials were excluded during data analysis. Each participant was tested with three runs, each lasting 7 min 42 sec.

To define the LOC and the inferior IPS ROIs, the same localizer experiment used in Xu and Chun (2006) was conducted here. Participants viewed blocks of object and noise images (both subtended $11.8^\circ \times 11.8^\circ$). The object images were the set size 6 displays used in the superior IPS localizer experiment. Each block lasted 16 sec and

contained 20 images, with each image appearing for 500 msec and followed by a 300 msec blank delay. To engage participants' attention to the displays, they were asked to detect a slight spatial jitter, which occurred randomly once in every 10 images. Each run contained eight object blocks and eight noise image blocks. Each participant was tested with two runs, each lasting 4 min and 40 sec.

fMRI Methods

fMRI data were acquired from a Siemens Tim Trio 3T scanner at the Harvard Center for Brain Science in Cambridge, MA. Participants viewed images back projected onto a screen at the rear of the scanner bore through an angled mirror mounted on the head coil. All experiments were controlled by an Apple MacBook Pro running Matlab with Psychtoolbox extensions (Brainard, 1997). Anatomical images were acquired using standard protocols. For both the localizer runs and the main experimental runs, 24 slices of 5 mm thick (3 mm × 3 mm in plane, 0 mm skip) parallel to the AC–PC line were acquired using a gradient-echo pulse sequence (echo time = 25 msec, flip angle = 90°, matrix = 64 × 64). In the main VSTM experiment and the superior IPS localizer runs, repetition time (TR) of 1.5 sec was used, and in the inferior IPS localizer runs, TR of 2.0 sec was used.

Data Analysis

Behavioral VSTM capacity for each set size was measured using Cowan's K formula, which estimates the number of items retained in VSTM while controlling for correct guesses ($K = [\text{hit rate} + \text{correct rejection rate} - 1] \times N$, wherein K is the number of items encoded in VSTM and N is the set size; see Cowan, 2001 for details).

fMRI data were analyzed with BrainVoyager QX 2.1 (www.brainvoyager.com). 3-D motion correction, slice acquisition time correction, linear trend removal, and Talairach space transformation were conducted during data preprocessing (Talairach & Tournoux, 1988).

To define the superior IPS ROI in each participant, as was done previously (Xu & Chun, 2006; Todd & Marois, 2004), fMRI data from the superior IPS localizer experiment were analyzed using multiple regressions with the regression coefficient for each VSTM set size weighted by that participant's behavioral VSTM capacity for that set size. The superior IPS was defined as voxels showing a significant activation in the regression analysis (false discovery rate $q < .05$, corrected for serial correlation) and whose Talairach coordinates matched those reported in Todd and Marois (2004). The LOC and inferior IPS ROIs were defined as voxels showing higher activations to the shape than to the noise displays (false discovery rate $q < .05$, corrected for serial correlation) in lateral occipital cortex and IPS, respectively. Example superior IPS, inferior IPS, and LOC ROIs are shown in Figure 2.

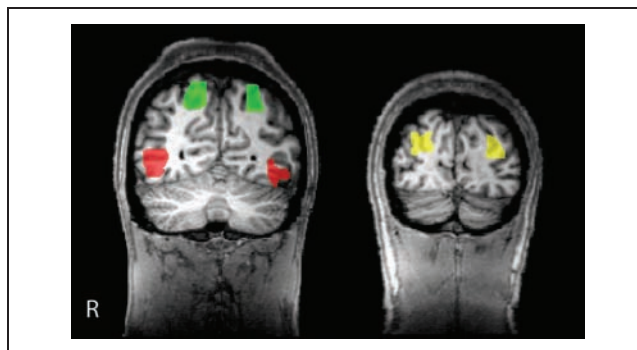


Figure 2. The superior IPS (green), the inferior IPS (yellow), and the LOC (red) ROIs from one example participant.

To examine responses from the main experiment, time courses from each participant in the main experiment were extracted from the three ROIs defined above. These time courses were converted to percent signal change for each stimulus condition by subtracting the corresponding value for the fixation trials and then dividing by that value (see Xu & Chun, 2006; Todd & Marois, 2004; Kourtzi & Kanwisher, 2000). To capture VSTM encoding-related peak responses in each participant and to account for temporal variability of fMRI peak responses among the different participants, VSTM encoding-related peak responses from all participants were aligned to the 9th second (6th TR) from the start of the trial. This anchor point was chosen based on responses from the majority of the participants. Time course data either remained the same or was shifted forward or backward by 1.5 sec (1 TR) during this alignment process. To ensure that baseline fMRI response differences before the onset of the VSTM shape display would not contribute to peak fMRI response amplitude estimates, we calculated baseline response drift by averaging the responses from the first 6 sec of each trial and then subtracted this drift from each point of the time course. This was done separately for each participant for each stimulus condition of each ROI.

Results

Behavioral Results

The capacity of VSTM was estimated using Cowan's K formula (Cowan, 2001). The mean K values for the four stimulus conditions were 0.89 ± 0.04 (1T), 0.90 ± 0.04 (1T + 3D), 1.90 ± 0.35 (4T), and 1.79 ± 0.3 (4T + 3D). A two-way repeated-measures ANOVA with Target Number (1 vs. 4) and Distractor Number (0 vs. 3) revealed a main effect of Target Number, $F(1, 11) = 11.063$, $p = .007$, showing that more information could be retained in VSTM from four targets than from one target. No other main effects or interactions reached significance ($F_s < 1$, $p_s > .57$).

RTs for the four stimulus conditions were 823 ± 49 msec (1T), 828 ± 36 msec (1T + 3D), 1003 ± 47 msec (4T), and

965 ± 45 msec (4T + 3D) respectively. Similar to the K measures, a two-way repeated-measures ANOVA with Target Number and Distractor Number revealed a main effect of Target Number, $F(1, 11) = 37.31, p < .001$, and a marginally significant interaction between Target Number and Distractor Number, $F(1, 11) = 3.4, p = .092$. No other main effect reached significance ($F < 1, p > .37$).

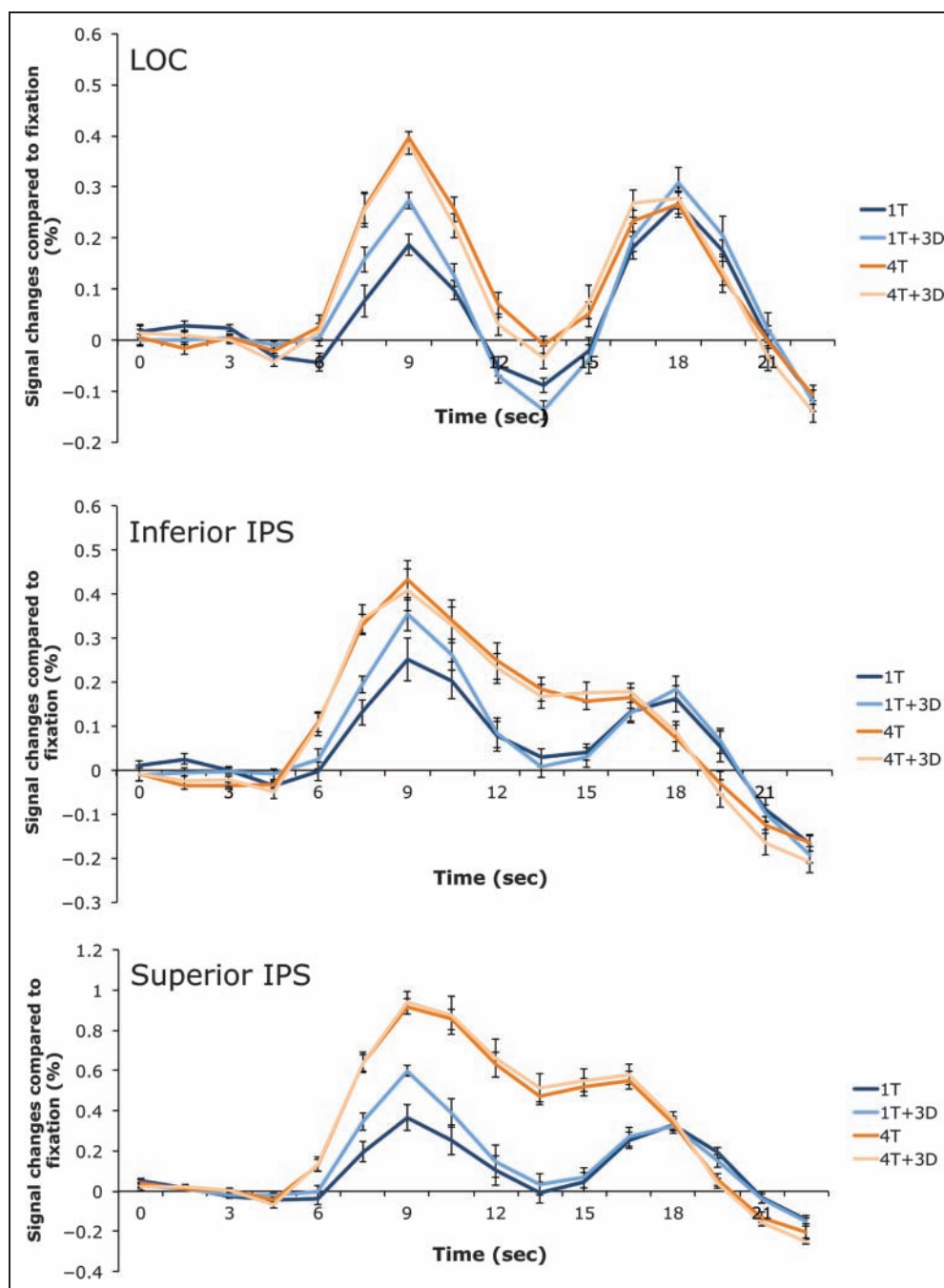
fMRI Results

fMRI responses from the main VSTM task were extracted from independently localized LOC, inferior IPS, and supe-

rior IPS ROIs. Percent signal change compared with fixation was calculated for each time point and the final time courses were plotted in Figure 3. These time courses showed two peaks, corresponding to the encoding of the initial shape display and the shape probe, respectively.

VSTM encoding-related activities. To examine VSTM encoding related activities, we analyzed the first fMRI peak responses at the 9th sec (6th TR) in the three ROIs. The effect of target number was present in superior IPS and LOC ($F_s > 25.47, p_s < .001$), but not in inferior IPS ($F < 1.93, p > .19$). The effect of distractor number was

Figure 3. fMRI responses from LOC (top), inferior IPS (middle), and superior IPS (bottom) in Experiment 1. All three brain regions showed a similar response pattern. The presence of distractors only increased fMRI responses during the encoding period when the target encoding load was low. No distractor effect was present during the delay period. Blue line, one target; light blue line, one target and three distractors; orange line, four targets; and light orange line, four targets and three distractors. Error bars indicate within-subject standard errors.



present in superior IPS, $F(1, 11) = 10.001, p = .009$, but not in the other two brain regions ($Fs < 2.41, ps > .148$). Importantly, all three brain regions showed a significant interaction between target number and distractor number ($Fs > 13.27, ps < .004$), indicating that distractor encoding was greater when one target than when four targets had to be encoded. Confirming this last result, in pairwise comparisons, in all three ROIs, significant differences were observed between 1T and 1T + 3D conditions ($Fs > 3, ps < .05$), but not between 4T and 4T + 3D conditions ($Fs < 1, ps > .58$). These results showed that distractor processing in inferior and superior IPS depended on target encoding load and only occurred at the low task load. Given that inferior and superior IPS have been proposed to be involved in object individuation and identification, respectively (Xu, 2007, 2009; Xu & Chun, 2006, 2009), these results suggest that distractor processing impacts both stages of object processing and is load dependent.

We also compared the difference between the 1T + 3D and 4T conditions in which the same total number of items were presented but target number differed. Interestingly, the difference between these two conditions was not significant in inferior IPS ($F < 1, p = .37$) but reached significance in both superior IPS and LOC ($Fs > 4.67, ps < .01$). In fact, the difference between these two conditions was greater in superior IPS than in inferior IPS ($F = 2.33, p = .039$). This may explain why we failed to obtain a main effect of target number in inferior IPS.

These results indicate that, when distractors were encoded under low target load, they were not differentiated from targets in inferior IPS that supports object individuation; the difference between targets and distractors only emerged in superior IPS and LOC that support object identification. This is consistent with the predictions of the neural object file theory proposed by Xu and Chun (2009). They argued that only object location information is predominantly encoded during object individuation and that detailed object feature information becomes available later during object identification-related processing (see also Xu, 2009).

VSTM delay-related activities. To examine VSTM maintenance-related activities, we analyzed fMRI responses at the 13.5th sec (9th TR) when responses reached a minimum before they started to rise again with the presentation of the probe display. During this delay period, a main effect of Target Number was observed in all three ROIs ($Fs > 10.24, ps < .01$), showing that four-target conditions elicited higher responses than one-target conditions. A main effect of Distractor Number was observed in LOC, $F(1, 11) = 6.796, p = .024$, showing a lower response for distractor present than for distractor absent conditions. Critically, there was no interaction between Target Number and Distractor Number in all three ROIs ($Fs < 1, ps > .5$). These results indicated that distractors either had no impact on target processing or they were completely sup-

pressed during the delay period. Either way, distractor processing did not depend on target processing load.

Discussion

By examining the impact of distractors on object individuation and identification during VSTM encoding and delay periods, here we observed neural encoding of distractors during both object individuation and identification when the target encoding load was low. The encoding of distractors under low load is consistent with the predictions of the load theory (Lavie, 2005; Lavie & Tsai, 1994).

Such load-dependent distractor response in inferior and superior IPS and LOC distinguishes them from pure stimulus-driven retinotopic visual regions. This is because, although almost twice the area was stimulated when four targets were presented with three distractors than when they were presented alone, we failed to observe any increase in response amplitude in these three brain regions.

When Xu (2010) examined the encoding to two features from the same object, she found that object-based encoding of task-irrelevant distractor features only occurred when the demand to encode the task-relevant target features was low. Because target and distractor features appeared on the same object and at the same location in Xu (2010), it might have been difficult to suppress the processing of distractor features. However, the present experiment showed that, even when targets and distractors appeared in different spatial locations and in different colors, distractor processing still could not be suppressed at low task encoding load. This indicates that the encoding of distractors at low task load may be automatic and obligatory.

Meanwhile, the present experiment showed that the neural response for distractors was short lasting and quickly decayed when no attempt was made to sustain it during the subsequent delay period. This is consistent with Xu (2010), which showed a similar response pattern for task-irrelevant features during object-based feature encoding. Thus, although the neural encoding of distractors at low target load may be initially automatic and obligatory, participants can exert control over what is retained for a prolonged period.

EXPERIMENT 2

In the Posner cuing paradigm (Posner, 1980), participants can better detect targets present at the cued than at the uncued spatial locations. This shows that the deployment of spatial attention can prioritize the processing of visual information at specific locations. Can such top-down attentional control suppress the processing of task-irrelevant distractors during target object individuation and identification? It is possible that with spatial cuing, neural encoding of distractors at low task load can be completely suppressed. It is equally likely, however, that, although the processing of distractors is attenuated, it cannot be

completely suppressed and that different amount of suppression may occur during target object individuation and identification. In this experiment, to understand how automatic and obligatory it is to encode task-irrelevant distractors under low load, we precued the locations of the targets before target onset and tested whether distractor encoding could be suppressed by top-down attention. Given that Experiment 1 showed that the presence of distractors had no impact on VSTM maintenance and retrieval-related activities (see Figure 3), to streamline our design, instead of using a 8.3-sec delay period, here we used a 1-sec delay period.

Methods

Participants

Nine new participants (seven women) were recruited from the Harvard University community (mean age = 28.33 years, $SD = 4.52$ years) with informed consent, which was approved by the institutional review board of Harvard University. All of them were right-handed and had normal or corrected-to-normal visual acuity. One additional participant was tested but excluded from further analyses due to excessive amount of head motion.

Design

The main VSTM experiment was identical to Experiment 1 except for the following. We shortened the delay period to 1000 msec, as the focus of this experiment was on distractor encoding. We also removed the verbal rehearsal load, as VSTM task performance with a short delay period has been shown to be unaffected whether verbal rehearsal is imposed or not (e.g., Luck & Vogel, 1997). In the valid cue trials, we cued target locations by rapidly flashing small dots twice at the target location before target onset. The neutral cue trials were similar to the valid cue trials, except that all eight locations where targets and distractors could possibly appear were cued by the flashing dots. To maximize the effect of cuing, valid and neutral cue trials were shown in different runs, with half of the participants tested with the valid cue trials before the neutral cue trials and the other half had the reverse order of testing. The exact timing of a trial was as follows: first precue (125 msec), a fixation period (125 msec), second precue (125 msec), a fixation period (625 msec), a sample display (200 msec), a delay period (1000 msec), a test shape display (1800 msec), and a feedback (2000 msec). Note that the 1000-msec interval between the initial onset of the cue and the onset of the stimulus display was the same as that used in Posner (1980). The participants were instructed to maintain fixation at the center fixation dot and covertly pay attention to the cued locations. Other aspects of this experiment were identical to those of Experiment 1.

Data Analyses

Because each trial lasted 6 sec with a 1-sec delay period, only one fMRI response peak was observed, reflecting the summed fMRI responses from VSTM encoding, maintenance, and retrieval. As such, instead of presenting data from each time point as we did in Experiment 1, only peak responses were extracted and included in further statistical analyses. All other aspects of data analyses were identical to that of Experiment 1.

Results

Behavioral Results

K values were 0.97 ± 0.015 (1T), 0.98 ± 0.01 (1T + 3D), 3.2 ± 0.16 (4T), and 2.71 ± 0.27 (4T + 3D) for neutral cue trials and 0.97 ± 0.015 (1T), 0.97 ± 0.015 (1T + 3D), 2.9 ± 0.15 (4T), and 2.93 ± 0.17 (4T + 3D) for valid cue trials. A three-way repeated measures ANOVA with cue type (neutral vs. valid), target number (1 vs. 4), and distractor number (0 vs. 3) was conducted. Main effect of target number was significant, $F(1, 8) = 220.44$, $p < .001$, showing that more information was stored in VSTM when target number was 4 than 1. No other main effects or interactions reached significance ($ps > .16$).

RTs were 496.6 ± 31.8 msec (1T), 481.4 ± 28.4 msec (1T + 3D), 753.6 ± 48.2 msec (4T), and 760.5 ± 44.2 msec (4T + 3D) for neutral cue trials, and 520.7 ± 26.7 msec (1T), 501.6 ± 24.9 msec (1T + 3D), 763 ± 48.4 msec (4T), and 776.1 ± 44.1 msec (4T + 3D) for valid cue trials. A three-way ANOVA with cue type, target number, and distractor number revealed a main effect of target number, $F(1, 8) = 131.08$, $p < .001$, showing that RT was slower when more targets had to be encoded and retrieved for comparison. There was also an interaction between target number and distractor number, $F(1, 8) = 5.521$, $p = .047$, indicating that RT difference between one- and four-target trials were larger when there were three than zero distractor. This is likely associated with the greater effort needed to filter out distractors at high than at low target encoding load. No other main effects or interactions reached significance ($ps > .36$).

fMRI Results

Although the peak fMRI responses examined here reflected the summed fMRI responses from VSTM encoding, maintenance, and retrieval periods, given that Experiment 1 showed that the presence of distractors had no impact on maintenance and retrieval related activities (see Figure 3), any distractor effect we obtained here could only come from encoding related activities.

In all three ROIs, as can be seen in Figure 4, there were a main effect of Targets, a main effect of Distractors, and an interaction between the two (all F s > 9.24 , $ps < .05$). This replicated our findings from Experiment 1 and showed that the presence of distractors significantly impacted target processing in a load-dependent manner.

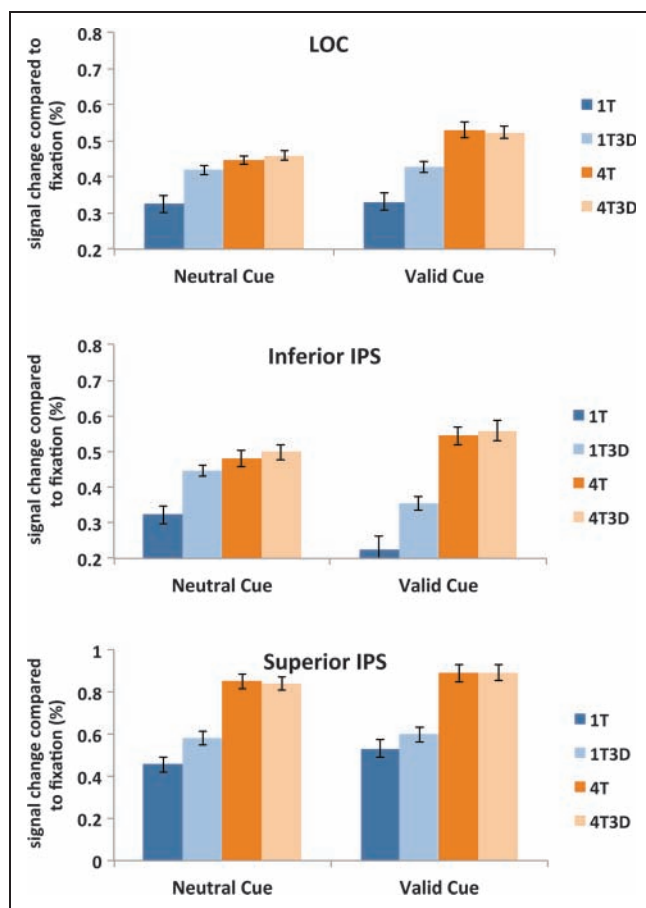


Figure 4. fMRI responses from LOC (top), inferior IPS (middle), and superior IPS (bottom) in Experiment 2. As in Experiment 1, the presence of distractors only increased fMRI responses when the target encoding load was low. Critically, although spatial cuing did not completely remove distractor processing, it did significantly attenuate distractor processing during target object processing in superior IPS. Blue bar, one target; light blue bar, one target and three distractors; orange bar, four targets; and light orange bar, four targets and three distractors. Error bars indicate within-subject standard errors.

Of main interest was the effect of spatial cuing. Out of the three ROIs, only the superior IPS showed a significant three-way interaction of Cue Type, Target Number, and Distractor Number ($F(1, 8) = 5.827, p = .042$; for inferior IPS and LOC, $F_s < 1, p_s > .6$). Detailed comparisons revealed that, in superior IPS, under low target encoding load, although the effect of distractor was still present in both the valid cue and the neutral cue conditions ($F = 2.74, p = .025$; and $F = 5.77, p < .001$, respectively), distractor processing was significantly attenuated with spatial cuing, resulting in a significant interaction between cue type and distractor number in the one-target conditions, $F(1, 7) = 18.167, p = .003$. Such an interaction, however, was absent in the four-target conditions ($F < 1, p > .63$). This pattern of response was found in every single one of our participants. Comparing directly across the three ROIs, there was a marginally significant interaction between the effect of cuing under low load and brain region, $F(2, 16) = 3.27, p = .064$, showing that

the effect of cuing under low load was stronger in superior IPS than in the other two brain regions.

In inferior IPS and LOC, there was an interaction between cue type and target number ($F_s > 16.8, p_s < .01$), showing that the difference between one- and four-target conditions was greater in the valid than in the neutral cue conditions. This could be due to differences in cue-related encoding, as in the valid cue trials, one cue and four cues were shown for the one- and four-target conditions, respectively; whereas in the neutral cue trials, eight cues were always shown regardless of the target encoding load. It is also possible that this interaction between cue type and target number was the result of more efficient allocation of resources with target cuing, such that less resources were allocated to the one-target conditions in the valid than in the neutral cue conditions and more resources were allocated to the four-target conditions in the valid than in the neutral cue conditions.

Discussion

This experiment replicated the main findings of Experiment 1 and showed that distractor processing only occurred under low target encoding load. Importantly, this experiment indicated that, although spatial cuing could not completely remove distractor processing, it could significantly attenuate distractor processing in superior IPS. This is consistent with a previous finding showing that superior IPS is mainly involved in processing what is most task relevant (Xu, 2010).

Unlike Posner (1980), here we did not observe any behavioral cuing benefit. This is likely due to the fact that our VSTM paradigm is not configured to produce the behavioral cuing benefit. In Posner's study, participants made speeded detection for the appearance of the cued target. In the present experiment, this was not measured. Rather, behavioral accuracy and RT mainly reflected responses to the shape probe 1 sec after the presentation of the target shapes. Nevertheless, the effect of cuing did impact distractor processing in superior IPS, showing that in this case fMRI measures could be more sensitive and informative than behavioral measures.

In our experiment, we used a fixed time interval between the initial onset of the cue and the onset of the stimulus display. It would be worth manipulating this interval in future studies to see whether distractor processing is modulated by this interval during target object individuation and identification. In any event, the 1000-msec cuing interval used in this experiment, which was the same as that used in Posner (1980), clearly illustrates the feasibility of using spatial cuing to prioritize the processing of targets among distractors.

GENERAL DISCUSSION

In this study, we investigated the processing of task-irrelevant information during visual object individuation

and identification by examining the neural substrates mediating these processes. We asked participants to encode in VSTM target object shapes among distractor object shapes appearing at different spatial locations and in different colors and examined fMRI responses from parietal and occipital regions. In Experiment 1, we found that distractor processing depended on the availability of processing resources. Specifically, only when the demand to encode target shapes was low did the presence of distractors increase neural responses in inferior IPS, LOC, and superior IPS. Given the involvement of these brain regions in object individuation and identification (e.g., Xu & Chun, 2009), these results suggest that, under low target encoding load, distractors were individuated and encoded. However, neural responses for distractors were short-lived as they were only present during the VSTM encoding period but not during the subsequent VSTM delay period. In Experiment 2, we examined whether distractor encoding under low task load could be suppressed if spatial attention was deployed ahead of the time to the target locations. Precuing target locations decreased distractor processing under low task load in superior IPS but not in inferior IPS or LOC. Thus, although distractor processing under low task load is obligatory and automatic during object individuation in inferior IPS and object encoding in LOC, it can be attenuated during object encoding in superior IPS with precuing.

Consistent with this result, Xu (2010) reported that superior IPS encoded only task-relevant features regardless of the target encoding load whereas LOC encoded task-irrelevant information at low load. Likewise, task-dependent encoding in parietal regions has also been reported in neurophysiology studies (Freedman & Assad, 2006; Toth & Assad, 2002). Thus, although distractor processing was not suppressed in superior IPS in Experiment 1, with the help of precuing, this brain region can exhibit some degrees of task-dependent responses in Experiment 2.

Results of this study, together with previous studies showing the impact of perceptual and working memory load on distractor processing in other visual tasks, support the perceptual load theory, which argues that the processing of distractors depends on the available resources and only occurs when the main task is relatively easy and does not consume all the available resources (Xu, 2010; Torralbo & Beck, 2008; Lavie, 2005; Lavie, Hirst, De Fockert, & Viding, 2004; Pinski, Doniger, & Kastner, 2004; Yi et al., 2004; Lavie & Tsai, 1994). Meanwhile, this study also identifies situations in which distractor processing under low task load may be suppressed (i.e., during the VSTM delay period) or substantially attenuated (i.e., with spatial cuing during object encoding in superior IPS).

Because distractor suppression related neuronal activities could also increase fMRI responses, one may argue that an increased fMRI response at low task load could reflect distractor suppression, rather than encoding. This, however, is unlikely the case due to the following two reasons. First, although distractor suppression was more

critical at high task load when participants needed to dedicate all their encoding resources to targets, we did not see an increased fMRI response to distractor processing at high task load in both Experiments 1 and 2. Second, with spatial cuing in Experiment 2, more suppression would be applied to distractors, and yet we observed an attenuated fMRI response to distractor processing at low task load in superior IPS and no response to distractor processing at high task load in all three ROIs. Thus, the distractor-related fMRI responses reported here reflect distractor encoding and not suppression.

Although this study showed that distractors could be individuated and encoded when the target encoding load was low, it is unknown whether target and distractor shapes were encoded with the same precision. When they are task relevant, shapes need to be encoded in sufficient resolution to support later memory recognition; when they are task irrelevant, however, shapes may not be encoded in such fine resolution. Recent studies using multivoxel pattern analysis have been able to decode visual information representation in a brain region by examining fMRI voxel response patterns (e.g., Norman, Polyn, Detre, & Haxby, 2006; Cox & Savoy, 2003; Haxby et al., 2001). Further research using the multivoxel pattern analysis approach may inform us of the exact nature of distractor shape representation during visual object individuation and identification.

In summary, the current study showed that, under low target encoding load, distractors elicited significant neural responses across a number of brain regions previously shown to be involved in visual object individuation and identification. This suggests that distractors are individuated and encoded at load target encoding load. However, such neural responses for distractors were short-lived as they were only present during the VSTM encoding but not the delay period. Although distractor processing was obligatory and automatic at low task load, with spatial cuing, it could be attenuated during object encoding in superior IPS.

Acknowledgments

We thank Aaron Glick and Sonia Poltoratski for their assistance in this study. This research was supported by NSF grant 0855112 to Y. X.

Reprint requests should be sent to Su Keun Jeong, Psychology Department, 33 Kirkland Street, Harvard University, Cambridge, MA 02138, or via e-mail: skjeong@fas.harvard.edu.

REFERENCES

- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.
- Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia, 6*, 255–265.

- Corkin, S. (2002). What's new with the amnesic patient H.M.? *Nature Reviews Neuroscience*, *3*, 153–160.
- Corkin, S., Amaral, D. G., González, R. G., Johnson, K. A., & Hyman, B. T. (1997). H. M.'s medial temporal lobe lesion: Findings from magnetic resonance imaging. *Journal of Neuroscience*, *17*, 3964–3979.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*, 87–114; discussion 114–185.
- Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) “brain reading”: Detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*, *19*, 261–270.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*, 80–90.
- Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature*, *443*, 85–88.
- Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, *14*, 477–485.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*, 3310–3318.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*, 75–82.
- Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*, 339–354.
- Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*, 183–197.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*, 616–618.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *92*, 8135–8139.
- Milner, B., Corkin, S., & Teuber, H. L. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of H.M. *Neuropsychologia*, *6*, 215–234.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*, 56–60.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*, 424–430.
- Pinsk, M. A., Doniger, G. M., & Kastner, S. (2004). Push-pull mechanism of selective attention in human extrastriate cortex. *Journal of Neurophysiology*, *92*, 622–629.
- Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*, 3–25.
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*, 65–97.
- Pylyshyn, Z. (1994). Some primitive mechanisms of spatial attention. *Cognition*, *50*, 363–384.
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, and Psychiatry*, *20*, 11–21.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-D proportional system: An approach to cerebral imaging*. New York: Thieme.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*, 751–754.
- Torr-albo, A., & Beck, D. M. (2008). Perceptual load-induced selection as a result of local competitive interactions in visual cortex. *Psychological Science*, *19*, 1045–1050.
- Toth, L. J., & Assad, J. A. (2002). Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature*, *415*, 165–168.
- Ungerleider, L. G., Courtney, S. M., & Haxby, J. V. (1998). A neural system for human visual working memory. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 883–890.
- Xu, Y. (2007). The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. *Journal of Neuroscience*, *27*, 11676–11686.
- Xu, Y. (2008). Representing connected and disconnected shapes in human inferior intraparietal sulcus. *Neuroimage*, *40*, 1849–1856.
- Xu, Y. (2009). Distinctive neural mechanisms supporting visual object individuation and identification. *Journal of Cognitive Neuroscience*, *21*, 511–518.
- Xu, Y. (2010). The neural fate of task-irrelevant features in object-based processing. *Journal of Neuroscience*, *30*, 14020–14028.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*, 91–95.
- Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *104*, 18766–18771.
- Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*, 167–174.
- Yi, D.-J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience*, *7*, 992–996.