Visual Short-term Memory Capacity for Simple and Complex Objects

Roy Luria[1], Paola Sessa[1], Alex Gotler[2], Pierre Jolicœur[3],
and Roberto Dell'Acqua[1]

Abstract

■ Does the capacity of visual short-term memory (VSTM) depend on the complexity of the objects represented in memory? Although some previous findings indicated lower capacity for more complex stimuli, other results suggest that complexity effects arise during retrieval (due to errors in the comparison process with what is in memory) that is not related to storage limitations of VSTM, per se. We used ERPs to track neuronal activity specifically related to retention in VSTM by measuring the sustained posterior contralateral negativity during a change detection task (which required detecting if an item was changed between a memory and a test array). The sustained posterior contralateral negativity, during the retention interval, was larger for complex objects than for simple objects, suggesting that neurons mediating VSTM needed to work harder to maintain more complex objects. This, in turn, is consistent with the view that VSTM capacity depends on complexity. ■

INTRODUCTION

Our visual system perceives a rich and highly detailed environment. We extract important visual information from this environment and use it to guide our behavior. This visual information is stored in a temporary buffer known as visual short-term memory (VSTM). VSTM can only maintain a limited amount of information in an "on-line" state, ready to be accessed or manipulated. To study the capacity limitations of VSTM, we often use a change detection paradigm, in which a memory array consisting of a small set of "objects" (e.g., colored squares, tilted bars, and so on) is presented briefly. After a blank retention interval of about 1 sec, a test array is presented, and participants indicate whether the test and the memory arrays are identical or whether they differ in some way. Accuracy in this change detection task is often very high when one, two, or up to three objects are presented and then declines as more objects are added to the memory array. This method yields an estimate of the capacity of VSTM of about three or four items (e.g., Wheeler & Treisman, 2002; Luck & Vogel, 1997; Pashler, 1988; Sperling, 1960).

Luck and Vogel (1997) demonstrated that VSTM capacity for objects containing a single feature was equivalent to capacity for certain types of multifeatured objects. These authors presented simple and complex objects (e.g., objects that were composed of a single dimension vs. objects that were composed of features that varied in two or more dimensions, such as colored bars) and compared conditions in which just one feature was presented and could have been changed (i.e., just color or just orientation) to conditions in which colored bars were presented and could have been changed either in color or in orientation. Interestingly, their results indicated that performance was identical between these conditions. Namely, performance did not deteriorate when presenting colored bars that could have been changed in both features (either in color or in orientation), relative to conditions in which only one dimension was presented and could have been changed (i.e., just color or just orientation). Thus, they argued that VSTM capacity is determined by the number of integrated objects rather than by the number of individual features (see also Vogel, Woodman, & Luck, 2001).

It should be noted that not all the aspects of this study were replicated, especially regarding objects that have multiple values from a single dimension (i.e., integrated stimuli that are composed of two colors; see Delvenne & Bruyer, 2004; Wheeler & Treisman, 2002). For example, Wheeler and Treisman (2002, Experiment 2) have found reduced accuracy for bicolored than for single colored squares. This raises the possibility that only memory systems for different features are independent of each other (so that they can store up to four objects each) at least to some extent (Olson & Jiang, 2002).

Perhaps the strongest challenge to the claim that VSTM represents a fixed number of objects, regardless of their complexity, comes from studies purported to measure VSTM capacity directly for classes of stimuli that differ in their complexity. Alvarez and Cavanagh (2004) observed a monotonic decrease in VSTM capacity when object complexity (defined by search efficiency) was increased. For

[1]University of Padova, Italy, [2]Ben-Gurion University of the Negev, Israel, [3]Université de Montréal, Québec, Canada

instance, the capacity for colored squares was approximately four objects, but it was only two objects when the stimuli were random polygons. Moreover, by using objects that were difficult to categorize (i.e., ovals with varying aspect ratios and color mixtures) with new values in each trial, Olsson and Poom (2005) showed that VSTM capacity decreased to only one object. These studies suggest that VSTM capacity is sensitive to the quality and complexity of the stored information and is not determined solely by the number of objects (see also Eng, Chen, & Jiang, 2005).

A recent study by Awh, Barton, and Vogel (2007) highlighted an aspect of the change detection paradigm that may be crucial for interpreting this apparent discrepancy in VSTM capacity estimates. The change detection paradigm involves many different stages of processing other than encoding and storage in VSTM. According to Awh et al., events taking place during the test phase of the paradigm, when the memory representation is compared with the test display, are responsible for the lower accuracy when complex stimuli are involved. They argued that comparing the two arrays is more error prone when the objects are more complex due to high similarity between the memory and the test arrays. To support this claim, Awh et al. manipulated selectively the difficulty of the comparison process. They compared instances in which the changed items between the test and the memory belonged to different categories (e.g., a cube that was replaced by a Chinese character), thus reducing the similarity between the memory and the test array to instances in which the changed items belonged to the same category (e.g., a cube that was replaced by a different cube). The results indicated that when the changed items switched categories, VSTM capacity was estimated at three to four items even for more complex stimuli.

Awh et al. (2007) concluded that when the comparison process involves objects that are easy to discriminate, VSTM capacity is limited only by the number of items that can be maintained simultaneously (Luck & Vogel, 1997). However, when the stimuli in the memory and the test array are highly similar, performance is also limited by the comparison process, contaminating estimates of VSTM storage capacity based on accuracy in the change detection task. This argument is important because it may explain previous result that found lower capacity for complex stimuli in terms of a mechanism not directly related to VSTM. Namely, VSTM stores three to four objects regardless of their complexity during the retention interval, and performance is limited by a separate process (sensitive to stimulus similarity/complexity) that compares the information held in VSTM.

One way to track the moment-by-moment deployment of resources during VSTM maintenance without incurring in the problem described by Awh et al. (2007) is through the use of ERP time locked to the onset of the memory array. Using this technique, Klaver, Talsma, Wijers, Heinze, and Mulder (1999) found a negative slow wave at posterior

scalp electrodes contralateral to the visual field in which to-be-encoded information was presented, which was hypothesized to be related to the retention of visual stimuli (i.e., random polygons) in VSTM. McCollough, Machizawa, and Vogel (2007), Vogel, McCollough, and Machizawa (2005), and Vogel and Machizawa (2004) followed up on this work in several experiments, using a bilateral presentation version of a change detection paradigm, in which only a cued portion of the memory array, either the left side or the right side of fixation, had to be remembered for comparison with the test array. Vogel et al. also observed a sustained posterior negativity that was larger over the hemisphere contralateral to the cued portion of the memory array, which they labeled contralateral delay activity (CDA). The CDA activity persisted throughout the retention interval, and the amplitude increased as the number of items held in VSTM increased; that is, averaged CDA amplitude was the lowest when only one item was remembered and then increased progressively as the memory set size increased, reaching an asymptotic limit at each individual's estimated VSTM capacity. Incorrect trials were characterized by lower CDA amplitude relative to correct trials, which was hypothesized to be consistent with the idea that subjects retained less, and likely insufficient, information during incorrect trials relative to correct trials. Moreover, tracking the time course of this wave revealed that individuals with higher VSTM capacity were more efficient at encoding and representing in VSTM only relevant information relative to individuals with lower VSTM capacity, who appeared to encode and to maintain in VSTM also irrelevant information.

This sustained activity was further used on various occasions to track the contents of VSTM in tasks that required VSTM but did not involve the change detection paradigm (e.g., Jolicœur, Brisson, & Robitaille, 2008; Brisson & Jolicœur, 2007a; Dell'Acqua, Sessa, Jolicœur, & Robitaille, 2006; Jolicœur, Sessa, Dell'Acqua, & Robitaille, 2006). For example, using a dual task paradigm (in which two different stimuli are presented in rapid succession), Brisson and Jolicœur (2007a) demonstrated a delayed transfer into VSTM of visual information for the second of two stimuli as a result of central interference presumably created by cognitive operations required to perform the first task (see also Jolicœur, Dell'Acqua, et al., 2007; Brisson & Jolicœur, 2007b; Dell'Acqua et al., 2006; Robitaille & Jolicœur, 2006). Specifically, this study monitored the sustained posterior contralateral negativity (SPCN; equivalent to CDA) for the second task in the PRP paradigm and provided evidence that SPCN onset was delayed as the presentation time of the stimuli for both tasks was reduced. In addition, when the first task became more difficult, a similar delay in the SPCN was observed. These results nicely demonstrated that concurrent central processing interferes with the encoding of information into VSTM.

On the basis of the relatively large consensus that the SPCN observed during the retention of visual information

encoded from a lateralized stimulus is a marker for the moment-by-moment contents of VSTM, we designed the present series of experiments to monitor through the SPCN component how visual information is represented in VSTM for stimuli that differ in complexity.

In all experiments, subjects performed a change detection task with simple stimuli (colors) and complex stimuli (random polygons). If VSTM capacity is determined only by the number of items, SPCN amplitude should be equal for different classes of stimuli, as long as the number of objects is identical. A recent study by Woodman and Vogel (2008) showed that different classes of stimuli produce different SPCN amplitude that was additive with set size. Namely, the same SPCN slope was maintained for colors and orientation. Thus, finding an interaction between the set size and the different classes of stimuli would indicate that VSTM capacity is also sensitive to stimulus complexity. Moreover, finding evidence that the SPCN amplitude is larger for complex stimuli (with a small set size of objects) will demonstrate that complex objects consume more capacity relative to simple objects. This evidence could not be attributed to an error-prone comparison process because the SPCN tracks the capacity of VSTM during the retention interval, before the test array is presented. Therefore, we avoided the confounding factor identified by Awh et al. (2007).

EXPERIMENT 1

In Experiment 1, participants performed a change detection VSTM task in which we presented either colored squares or random polygons as stimuli. Figure 1A illustrates the sequence of events or trials with colored squares, and Figure 1B shows the set of polygons used in the experiment. To make it more difficult for participants to use verbal codes rather than VSTM, half of the participants performed the change detection task with a concurrent silent rehearsal task (i.e., they noiselessly rehearsed the names of two digits during the retention interval). The other half performed the task without concurrent silent rehearsal. Vogel et al. (2001) showed that this type of task causes a very large decrement in verbal memory and very little decrement in VSTM for colored squares. Although previous research using simple colored squares showed that verbal codes are not used in the present paradigm (i.e., Todd & Marois, 2004; Luck & Vogel, 1997), the use of verbal codes may still arise when two distinct categories are present, as in the present experiment. Interestingly, previous studies did not control for this factor (Eng et al., 2005; Olsson & Poom, 2005; Alvarez & Cavanagh, 2004).

Methods

In this section, we outline general aspects of the method that were common to all experiments and the modifications that distinguished each experiment from another.

Participants

In Experiment 1, 32 undergraduate students participated in a 2-hr session. Sixteen participants performed the experiment with a concurrent silent rehearsal task, and 16 participants performed the experiment without this concurrent

Figure 1. (A) The change detection paradigm. Each trial began with a fixation point (500 msec) followed by an arrow cue (presented for 400 msec) that indicated the relevant side for the up coming trial. Then, a memory display containing two to four objects, on each side of fixation, was presented for 100 msec, followed by a 900-msec retention period and then by a test array. Subjects judged whether the memory and the test array were identical or whether one object was different. (B) The set of the polygons used in the experiments.
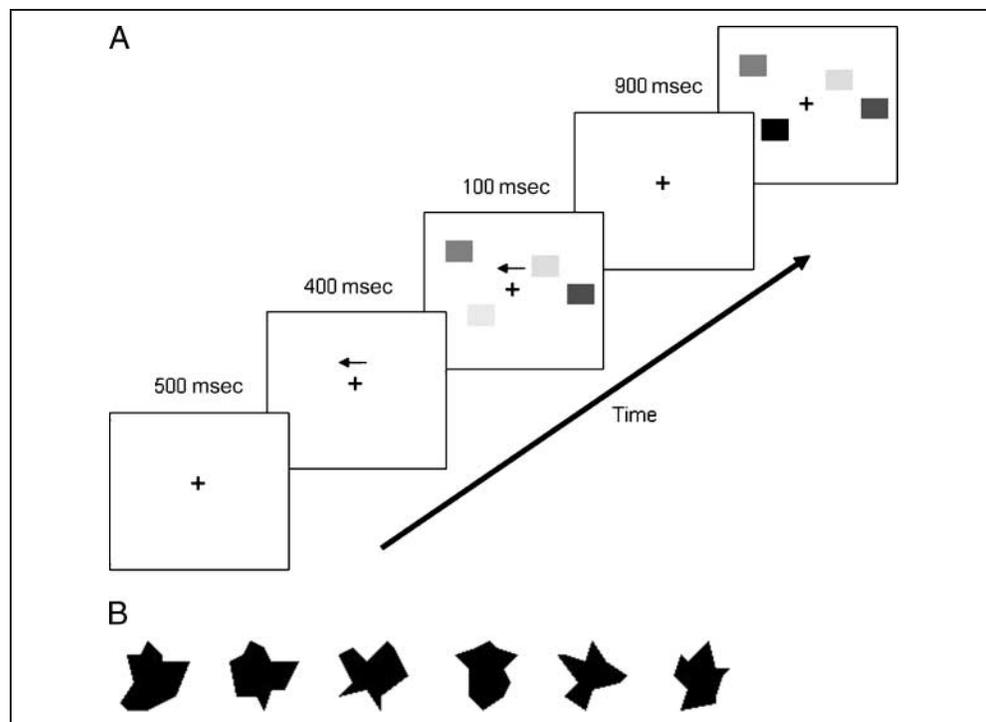
task. Nineteen undergraduate students participated in Experiment 2, 12 undergraduate students participated in Experiment 3, and 24 undergraduate students participated in Experiment 4. All participants reported no history of neurological problems and had normal or corrected-to-normal vision. Each experiment used a unique sample of participants.

Stimuli and Procedure

Visual stimuli were displayed on a gray background on a 17-in. cathode ray tube monitor controlled by a microcomputer running E-Prime software. In Experiment 1, the stimuli were either colored squares or random polygons. From a viewing distance of approximately 60 cm, each square subtended approximately $1.0°$ of visual angle in height and width, and each polygon subtended approximately $1.2° \times 1.2°$ of visual angle.

The exact stimuli were randomly selected at the beginning of each trial (from a pool of seven possibilities), with the restriction that any stimulus could appear no more than twice in an array (never on the same side). The colors were highly discriminable (red, blue, violet, green, yellow, black, and green). We used roughly the same random polygons as Alvarez and Cavanagh (2004), which were presented in black (see Figure 1B). Stimuli appeared in a $3.5° \times 7°$ rectangle (one in each side of fixations). Inside each rectangle, the exact positions of the stimuli were randomized on each trial, with the constraint that the distance between the upper-left corners of each stimulus would be more than $1.5°$ apart.

For the silent rehearsal group, each trial began with the presentation of two different digits (between 0 and 9, randomly determined). Participants were instructed to rehearse the digits silently until they emitted their response. After 500 msec, the digits were replaced by a fixation cross (subtending $1.4° \times 0.7°$) for 500 msec, followed by an arrow cue pointing to either the left or the right, which was presented for 400 msec, followed by the memory array. The memory array was composed of two, three, or four stimuli in each hemifield (randomly determined) and was presented for 100 msec, followed by a blank screen (except for the fixation cross which remained visible) for 900 msec, and then by the test array, as illustrated in Figure 1A. The test array was presented until a response was emitted. Participants were instructed to memorize only the stimuli presented on the side indicated by the arrow. For the group without the concurrent silent rehearsal, the design was identical, except that the two digits were not presented. On half of the trials, one item in the test array was different from the memory array (always on the memorized side). The new item was not already present in the memory array and was always from the same category (i.e., a color was replaced by a different color). It was presented in the same spatial position as the old item. On the other half of the trials, the memory and the test arrays were identical. Partici-

pants responded using the "F" and "J" keys (response keys on the second row from the bottom on a computer keyboard), using the index finger of the left and the right hand, respectively. The mapping between keys and responses was counterbalanced between participants.

Colors and polygons were presented in separate blocks. Participants started with a practice color block of 12 trials, then only the silent rehearsal group preformed another practice color block of 12 trials with the silent rehearsal task, followed by polygon practice block of 20 trials (for both groups). There were six experimental blocks, 120 trials each, that were ordered: color, polygon, polygon, color, color, and polygon. This order was chosen to average practice and other order effects across conditions.

In Experiment 2, the polygons were displayed in different colors, using the same set of colors as in Experiment 1. Participants were informed before each block which stimulus dimension (color or shape) was relevant for the upcoming block of trials. All participants performed the task with silent rehearsal secondary task as in Experiment 1.

Experiment 3 was identical to Experiment 2 (including the silent rehearsal task), except that the memory array presentation time was prolonged to 300 msec and that the interval between the offset of the arrow cue and the onset of the memory array was 200, 300, or 400 msec (randomly determined).

In Experiment 4, we presented in different blocks either different (low-similarity) colors (the same stimuli as in the previous experiments) or high-similarity colors that were all shades between green and blue (CIE x, y, luminance: 0.287, 0.216, 57.01 cd/m^2; 0.245, 0.393, 87.30 cd/m^2; 0.457, 0.402, 21.50 cd/m^2; 0.167, 0.152, 31.9 cd/m^2; 0.604, 0.340, 23.40 cd/m^2; 0.211, 0.244, 32.30 cd/m^2; and 0.269, 0.511, 63.60 cd/m^2). In each trial, two, four, or six stimuli (randomly determined) were presented on each side of fixation. Each colored square subtended approximately $0.5° \times 0.5°$ of visual angle. Participants were informed before each block, which condition (high or low similarity) would be presented. The silent rehearsal task was not used in this experiment. If subjects attempted to recode the colors verbally, they would have had a much easier task for colors in the low-similarity set, which all had obvious names. This strategy would be expected to reduce the amplitude of the SPCN for low-similarity trials relative to high-similarity trials. The reason is that instead of storing the information in VSTM, the subject would rely on verbal codes, a mechanism that is outside VSTM. This should result in less capacity demand for VSTM and thus to lower SPCN amplitudes. As we show below, however, the SPCN amplitudes did not differ across conditions.

EEG/ERP

EEG activity was recorded continuously with tin electrodes located at sites Fp1, Fp2, Fz, F3, F4, F7, F8, C3, C4, Cz, P3, P4, Pz, O1, O2, T7, T8, P7, and P8 (see Pivik

et al., 1993), referenced to the left earlobe. Horizontal EOG activity was recorded bipolarly from electrodes positioned on the outer canthi of both eyes. Vertical EOG activity was recorded bipolarly from two electrodes, above and below the left eye. Impedance at each electrode site was maintained below 5 kΩ. EEG, horizontal EOG, and vertical EOG activity was amplified, band-pass filtered using 0.01–80 Hz, and digitized at a sampling rate of 250 Hz. The EEG was algebraically re-referenced off-line to the average of the left and right earlobes and segmented into 1100-msec epochs starting from 100 msec before the memory array onset. Single trials with ocular artifacts (exceeding 200 μV) and with other artifacts (zero lines, transients, large fluctuations, and amplifier saturations) were excluded from analysis. Overall rejection rates were 12% in the color condition and 13% in the polygon condition (Experiment 1), 14% in the color condition and 17% in the polygon condition (Experiment 2), 12% in the color condition and 14% in the polygon condition (Experiment 3), and 14% in the different color condition and 13% in the similar color condition (Experiment 4). Only correct trials were included in the analysis. Separate average waveforms for each condition were then generated, and difference waves were constructed by subtracting the average activity recorded from the P7/P8 (and O1/O2) electrodes ipsilateral to the memorized array from the average activity recorded from P8/P7 (and O2/O1) electrodes contralateral to the memorized array.

Note that McCollough et al. (2007), Vogel et al. (2005), and Vogel and Machizawa (2004) averaged four different electrodes across lateral, occipital, and posterior parietal sites, whereas we measured the SPCN where it was the largest in our setup, namely, at P7/P8, to examine the SPCN where we had the highest signal-to-noise ratio. For completeness, we also report the results for O1/O2 electrodes (where the effects found at P7/P8 were present numerically, albeit less accentuated and not always statistically significant).

Results and Discussion

Behavioral Results

The accuracy data revealed a decrease in accuracy as set size increased, which was more pronounced for the random polygons than for the colors (see Table 1). An ANOVA on accuracy included the independent variables Condition (colors vs. polygons), Set Size (2, 3, or 4), and Group (with or without silent rehearsal). The ANOVA yielded main effects of Condition (0.93 for colors vs. 0.67 for polygons), $F(1,30) = 1109.62$, $p < .0001$, $MSE = 0.002$, and Set Size, $F(2,60) = 160.33$, $p < .0001$, $MSE = 0.001$. Group did not reach significance level or any other interaction involving this variable (all $Fs < 1$). The interaction between Condition and Set Size was also significant, $F(2,60) = 8.98$, $p < .002$, $MSE = 0.001$, reflecting a more pronounced drop in accuracy between Set Sizes 2 and 3

Table 1. Mean Accuracy (Proportion Correct) and Reaction Time (msec) in Experiments 1–4 for All Set Sizes and Stimulus Conditions

Set Size		Colors	Polygons
Experiment 1			
2	AC	0.97	0.75
	RT	742	1032
3	AC	0.93	0.65
	RT	803	953
4	AC	0.88	0.61
	RT	867	990
Experiment 2			
2	AC	0.95	0.70
	RT	734	876
3	AC	0.91	0.62
	RT	808	919
4	AC	0.86	0.59
	RT	873	955
Experiment 3			
2	AC	0.95	0.78
	RT	823	966
3	AC	0.89	0.67
	RT	929	981
4	AC	0.82	0.63
	RT	967	1032

Set Size		Low Similarity	High Similarity
Experiment 4			
2	AC	0.96	0.75
	RT	918	1077
4	AC	0.81	0.63
	RT	1054	1151
6	AC	0.68	0.58
	RT	1116	1202

AC = mean accuracy (%); RT = mean reaction time (msec).

for polygons than for colors, $F(1,30) = 21.56$, $p < .001$, $MSE = 0.001$.

Although our instructions specifically emphasized accuracy, we also analyzed RT to check for speed–accuracy trade-offs. The results are reported in Table 1. An ANOVA considering the same variables as those for the analyses

on accuracy yielded main effects of Condition, $F(1,30) = 17.95$, $p < .005$, $MSE = 94075.42$, indicating that responses in the color condition were 188 msec faster relative to the polygon condition, and Group, $F(1,30) = 6.85$, $p < .05$, $MSE = 254191.01$, reflecting the fact that RT was 190 msec faster with silent rehearsal. The results did not suggest speed-accuracy trade-offs.

Electrophysiology

P7/P8. Figure 2A shows the ipsilateral and the contralateral waveforms at P7/P8 and O1/O2 for the color squares condition for each level of memory set size, and Figure 2B shows the waveforms for the polygon condition.

Table 2 shows the mean SPCN amplitude in all conditions in a time window of 450–900 msec relative to the onset of the memory array. The mean SPCN amplitude was the lowest for two colors, intermediate for three colors, and highest for four colors, replicating previous findings (McCollough et al., 2007; Vogel et al., 2005; Vogel & Machizawa, 2004). Contralateral minus ipsilateral waveforms are shown in Figure 3 for the color condition and in Figure 4 for the polygon condition. The ANOVA for the mean amplitude of the SPCN (P7/P8 electrodes) within a time window of 450–900 msec postmemory array onset that included the same variables as the accuracy analysis yielded a main effect of Set Size, $F(2,60) = 3.38$, $p < .05$, $MSE = 0.67$, and an interaction between Condition and Set Size, $F(2,60) = 3.33$, $p < .05$, $MSE = 0.91$. This

Figure 2. Grand average waveforms contralateral and ipsilateral to the side indicated by the arrow at P7/P8 and O1/O2 electrode sites time locked to the onset of the memory array and collapsed across visual fields in Experiment 1 (for both rehearsal groups). (A) Waveforms for the colors condition. (B) Waveforms for the polygon condition. For display purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).
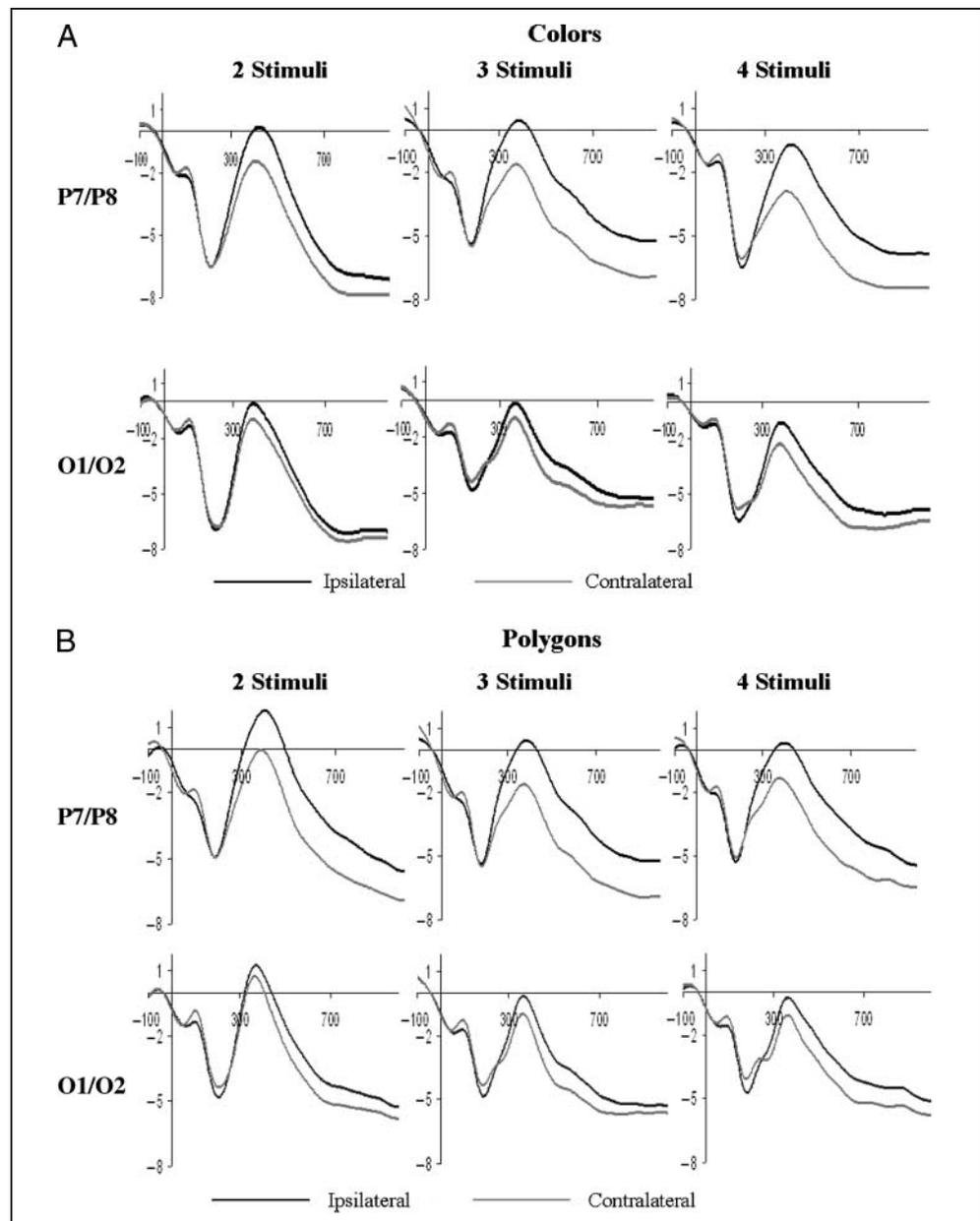
Table 2. Mean SPCN Amplitudes (for P7/P8 Electrodes) in a Window of 450–900 msec from the Onset of the Memory Array (μV) in Experiments 1–4 for All Set Sizes and Stimulus Conditions

Set Size	Colors	Polygons
Experiment 1		
2	−1.22	−1.83
3	−1.72	−1.98
4	−1.99	−1.73
Experiment 2		
2	−1.25	−2.22
3	−1.72	−1.76
4	−2.18	−1.38
Experiment 3		
2	−1.04	−1.92
3	−1.64	−1.60
4	−1.92	−1.58

Set Size	Low Similarity	High Similarity
Experiment 4		
2	−0.95	−0.85
4	−1.11	−1.11
6	−1.40	−1.11
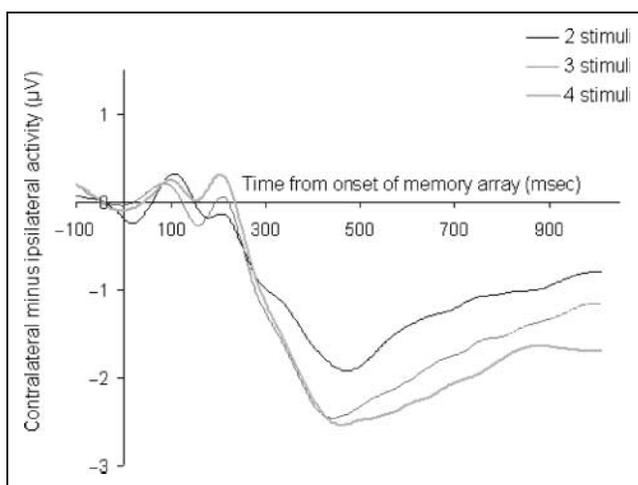
Figure 3. Grand average ERP waveforms time locked to the onset of the memory array, averaged across P7/P8 electrodes in Experiment 1, color condition for both rehearsal groups. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).
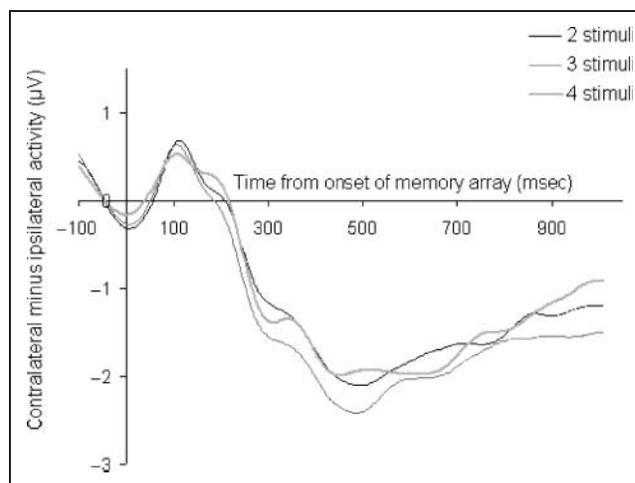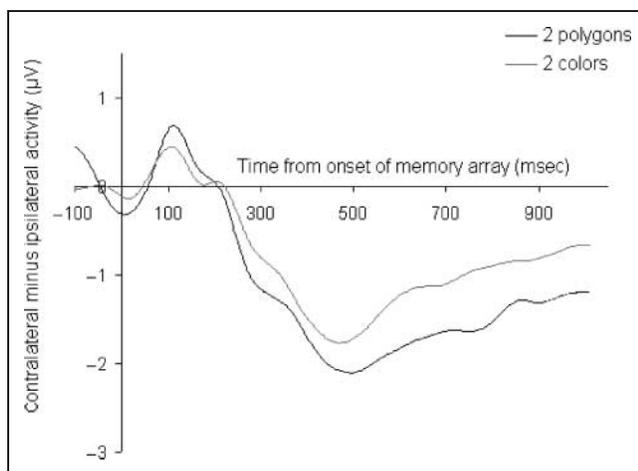
Figure 4. Grand average ERP waveforms time locked to the onset of the memory array, averaged across P7/P8 electrodes in Experiment 1, polygon condition for both rehearsal groups. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

interaction reflected the increase in SPCN amplitude (i.e., a greater contralateral negativity) as the set size increased for colors, $F(1,30) = 8.11, p < .01, MSE = 1.16$, but not for polygons, $F < 1$. The pattern of means suggests that SPCN amplitude for polygons was already at maximum for Set Size 2. In addition, we analyzed the slope of the change in mean SPCN amplitude across set size for colors and polygons for each subject and submitted them to an ANOVA that considered condition (colors vs. polygons) as a within-subjects variable. The slope for the color condition was −0.39 μV/item, and the slope for the polygon condition was 0.05 μV/item, $F(1,30) = 4.18, p < .05, MSE = 1.43$.

Given that the SPCN amplitude did not increase beyond two objects when polygons were encoded into VSTM, we also performed a direct comparison between two colors and two polygons. As can be seen in Figure 5, the SPCN had a higher amplitude for polygons relative to colors, although the number of items was the same, $F(1,30) = 4.70, p < .05, MSE = 1.24$. Importantly, finding larger SPCN amplitude for two complex objects than for two simple objects, coupled with finding an interaction between set size and condition, suggests that neurons mediating VSTM needed to work harder to maintain more complex objects. This, in turn, is consistent with the view that VSTM capacity depends also on complexity and not only on the number of objects.

Finally, silent rehearsal had no significant effect on the SPCN amplitude, $F(1,30) = 1.80, p > .18$ (nor did any interactions including this variable approached significance, highest $F = 1.52, p > .22$ in all cases).

O1/O2. The ANOVA with the same independent variables and with O1/O2 SPCN amplitude as a dependent variable (see Table 3) yielded only a significant interaction

Figure 5. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 1, for two colors and two polygons for both rehearsal groups. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

between Condition and Set size, $F(2,60) = 5.77$, $p < .05$, $MSE = 0.43$, indicating that for colors, the SPCN amplitude decreased with increasing the Set Size, $F(1,30) = 7.16$, $p < .05$, $MSE = 0.31$. For polygons, the opposite trend was observed, $F(1,30) = 4.26$, $p < .05$, $MSE = 0.62$. The difference between two polygons and two colors was not significant, $F = 1.27$, $p > .26$. Thus, the overall pattern was the same as in P7/P8, but the difference for two items was smaller and not significant.

EXPERIMENT 2

The results of Experiment 1 are in line with the hypothesis that more complex stimuli take up more VSTM capacity, as reflected in the larger SPCN amplitude for polygons relative to colors (e.g., Jolicœur, Brisson, et al., 2008; Jolicœur, Dell'Acqua, et al., 2007; Perron et al., 2009; Vogel & Machizawa, 2004). However, in Experiment 1, polygons and colors differed physically under a number of aspects, and one account that may be offered is that it was this feature of the design of Experiment 1, and not VSTM maintenance load, to produce the observed results. Experiment 2 was designed to control for this potential confound, by presenting stimuli that were physically identical (i.e., colored polygons) and by asking participants to ignore the polygons' shape and to attend only to their colors on some blocks of trials or vice versa in the other blocks of trials. If it was the physical dissimilarity of the stimuli used in the two conditions of Experiment 1 that caused differences in ERPs, then the results of Experiment 2 should show a reduction in the observed differences between SPCN amplitudes for colors and polygons.

Results and Discussion

Behavioral Results

As in Experiment 1, accuracy was lower for random polygons than for colors, and it declined as set size increased (see Table 1). The ANOVA included the same variables as in Experiment 1 (without the concurrent silent rehearsal variable). It yielded significant main effects of Condition, $F(1,18) = 730.57$, $p < .0001$, $MSE = 0.002$ (mean accuracy was better for colors, 0.91, than for polygons, 0.64), and Set Size, $F(2,36) = 74.60$, $p < .0001$, $MSE = 0.001$ (reflecting a decrease in accuracy as set size increased: 0.83, 0.77, and 0.73 for Set Sizes 2, 3, and 4, respectively). The interaction between Condition and Set Size was also significant, $F(2,36) = 4.07$, $p < .05$, $MSE = 0.0009$, indicating that the difference in accuracy between colors and polygons was smaller for Set Size 2 (0.25) relative to Set Size 3 (0.30), $F(1,18) = 7.86$, $p < .05$, $MSE = 0.001$.

Although our instructions specifically emphasized accuracy and mainly to rule out speed–accuracy trade-offs, we also analyzed RT (see Table 1). The ANOVA that

Table 3. Mean SPCN Amplitudes (for O1/O2 Electrodes) in a Window of 450–900 msec from the Onset of the Memory Array (μV) in Experiments 1–4 for All Set Sizes and Stimulus Conditions

Set Size	Colors	Polygons
Experiment 1		
2	−0.78	−0.98
3	−1.12	−0.87
4	−1.16	−0.57
Experiment 2		
2	−0.06	−0.48
3	−0.66	−0.68
4	−0.53	−0.53
Experiment 3		
2	−0.49	−1.05
3	−0.70	−0.99
4	−1.34	−0.99

| Set Size	Low Similarity	High Similarity
Experiment 4		
2	0.06	0.09
4	−0.12	−0.006
6	−0.32	−0.24

included the same variables as the analyses of the accuracy results yielded main effects of Condition, $F(1,18) = 12.89$, $p < .005$, $MSE = 27561.32$ (indicating that responses in the color condition were 112 msec faster relative to the polygon condition), and Set Size, $F(2,36) = 18.85$, $p < .001$, $MSE = 6020.22$ (reflecting an increase of 110 msec from Set Size 2 to Set Size 4). The results did not suggest speed-accuracy trade-offs.

Electrophysiology

P7/P8. The grand average subtraction waveforms (contralateral − ipsilateral) for each set size in the color condition are shown in Figure 6 (Table 2 presents the mean SPCN amplitude for all conditions). The amplitude of the SPCN was lowest for two colors, intermediate for three colors, and highest for four colors, replicating the results observed in Experiment 1 and in previous findings (Jolicœur, Brisson, et al., 2008; Jolicœur, Dell'Acqua, et al., 2007; Perron et al., 2009; McCollough et al., 2007; Vogel et al., 2005; Vogel & Machizawa, 2004). The grand average difference waveforms for the random polygon condition are presented in Figure 7, and they showed a markedly different pattern, as a function of set size, compared with the color condition—SPCN amplitude decreased with increasing set size rather than the opposite. The ANOVA on the SPCN amplitude, with a time window of 450–900 msec after onset of the memory array (with the same variables as Experiment 1) confirmed that the interaction between Condition and Set Size, $F(2,36) = 6.69$, $p < .05$, $MSE = 1.12$, was significant. The increase in the SPCN amplitude as the set size increased was significant for colors, $F(1,18) = 7.26$, $p < .05$, $MSE = 1.14$, but for polygons the amplitude with four items was low-
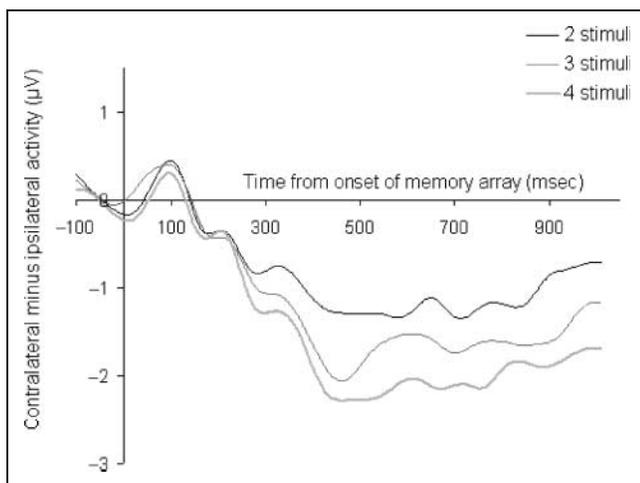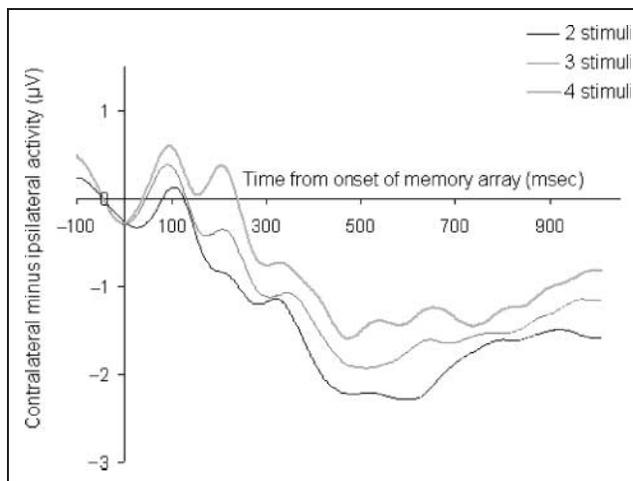
Figure 7. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 2, polygon condition. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

er relative to two items, a trend that was not significant, $F(1,18) = 2.84$, $p > .11$. One possible explanation for these divergent patterns of results is that in the most demanding conditions that likely surpass the maximum VSTM capacity, subjects choose to encode only part of the display (perhaps only in a subset of the trials).

In addition, we analyzed the slopes of the change in amplitude across set size for colors and polygons. The slope for the color condition was −0.47 μV/item and the slope for the polygons[1] was 0.48 μV/item. The difference between these slopes was significant, $F(1,18) = 10.38$, $p < .05$, $MSE = 1.45$, corroborating the interaction found in our analysis of mean amplitudes.

As in Experiment 1, we also compared the SPCN at Set Size 2 for colors and polygons (see Figure 8). The SPCN amplitude was higher for polygons relative to colors, although the number of items was the same, $F(1,18) = 4.42$, $p < .05$, $MSE = 2.04$.

O1/O2. The ANOVA with the same independent variables and with O1/O2 SPCN amplitude as a dependent variable (see Table 3) yielded only a significant main effect of Set Size, $F(2,36) = 3.55$, $p < .05$, $MSE = 0.35$, indicating that the SPCN amplitude decreased with increasing the Set Size, $F(1,30) = 4.96$, $p < .05$, $MSE = 0.51$. Although the interaction between Condition and Set Size was not significant, $F(2,36) = 1.74$, $p > .18$, $MSE = 0.01$, and inline with previous analysis, Set Size had a marginal significant effect for colors, $F(1,18) = 3.78$, $p = .06$, $MSE = 0.55$, but the same effect was far from being significant for polygons, $F < 1$. Although numerically present, the difference between two polygons and two colors was not significant, $F(1,18) = 1.58$, $p > .22$, $MSE = 1.03$. Thus, again the overall numerical pattern was similar as in P7/P8, but the interaction and the difference between two items were smaller and nonsignificant.

Figure 6. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 2, color condition. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).
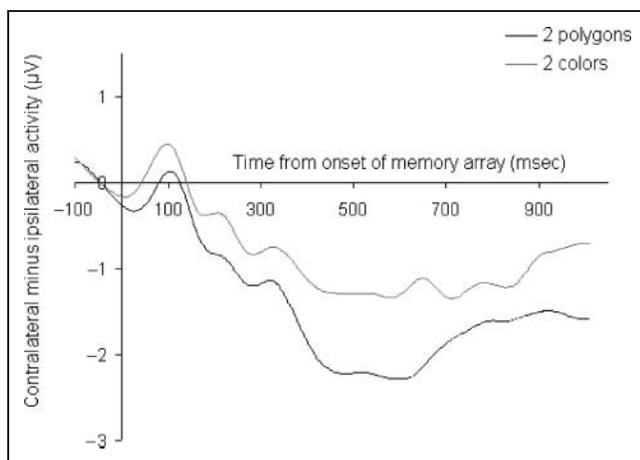
Figure 8. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 2, for two colors and two polygons. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

Overall, Experiment 2 replicated the most important aspects of results of Experiment 1 despite the use of physically identical stimuli (namely, colored polygons). The SPCN was larger when subjects encoded the shape of the polygons rather than their color, with the SPCN amplitude reflecting capacity saturation for only two objects under the instructions to encode the shape of the polygons in VSTM. Increasing the number of colors to be encoded increased the amplitude of the SPCN, but increasing the number of polygons did not further increase the amplitude of the SPCN.

EXPERIMENT 3

In Experiments 1 and 2, the memory array was exposed for 100 msec to decrease the likelihood of verbal recoding of the stimuli and to discourage eye movements toward the stimuli to be encoded (lateralized to the left or right visual field). In Experiment 3, we used a longer exposure duration to determine whether results from Experiments 1 and 2 may have been partly due to the short exposure duration used in these experiments. One might wonder, for example, if this brief presentation was a limiting factor when encoding information in VSTM rather than in VSTM maintenance per se. If this were so, one would expect this limiting factor to exert a greater effect for more complex stimuli, which could, perhaps, explain why the SPCN amplitude for four polygons was reduced relative to the two polygons condition (Figures 4–7). Perhaps, subjects found it difficult to encode all the information present in 100 msec displays and resorted to encode only a fraction of the information contained in the displays. Experiment 3 was designed to determine whether the very short exposure time of the memory array was critical to obtain the differences in SPCN across the color

and shape conditions in Experiments 1 and 2. Experiment 3 was practically identical to Experiment 2, with the important exception that the exposure time for the memory array was increased to 300 msec, providing subjects with substantially more time to encode the colored polygons presented in the memory array. If the SPCN amplitude difference documented so far for polygons reflected primarily a limitation at a stage of VSTM encoding, as opposed to VSTM maintenance, the results of Experiment 3 should show an attenuation of the differences observed in Experiments 1 and 2.

Results and Discussion

Behavioral Results

Accuracy was lower for random polygons than for colors, and it declined as set size increased (see Table 1). The ANOVA included the same variables as in Experiment 2. It yielded significant main effects of Condition, $F(1,11) = 24.52$, $p < .001$, $MSE = 0.02$ (mean accuracy was better for colors, 0.89, than for polygons, 0.69), and Set Size, $F(2,22) = 88.85$, $p < .0001$, $MSE = 0.001$ (reflecting a decrease in accuracy as set size increased: 0.87, 0.78, and 0.73 for Set Sizes 2, 3, and 4, respectively). To verify the effect of the prolonged presentation time, we performed another ANOVA including the results from Experiment 2 and the current one, with Experiment as an independent between-participants variable. Importantly, the interaction between Experiment and Condition was significant, $F(1,29) = 5.36$, $p < .05$, $MSE = 0.01$, indicating that for colors, the overall accuracy was the same between experiments ($F = 1.05$, $p > .31$). However, for polygons, the overall accuracy was higher in Experiment 3 relative to Experiment 2 (0.69 and 0.64, respectively), $F(1,29) = 7.89$, $p < .01$, $MSE = 0.008$. No other interactions involving Experiment were significant.

To rule out speed-accuracy trade-offs, we also analyzed RT (see Table 1). The ANOVA that included the same variables as for the analyses of the accuracy results yielded a main effect of Condition and Set Size, $F(2,22) = 9.96$, $p < .001$, $MSE = 6720.58$ (reflecting an increase of 105 msec from Set Size 2 to Set Size 4). The main effect of Condition was marginally significant, $F(1,11) = 3,78$, $p < .071$, $MSE = 35743.78$. The results did not suggest speed-accuracy trade-offs.

Electrophysiology

P7/P8. The grand average subtraction waveforms (contralateral − ipsilateral) for each set size in the color condition are shown in Figure 9, and the mean SPC amplitudes are listed in Table 2. The amplitude of the SPCN was lowest for two colors, intermediate for three colors, and highest for four colors, replicating the results observed in Experiments 1 and 2. The grand average difference waveforms for the polygon condition are presented in Figure 10, and they
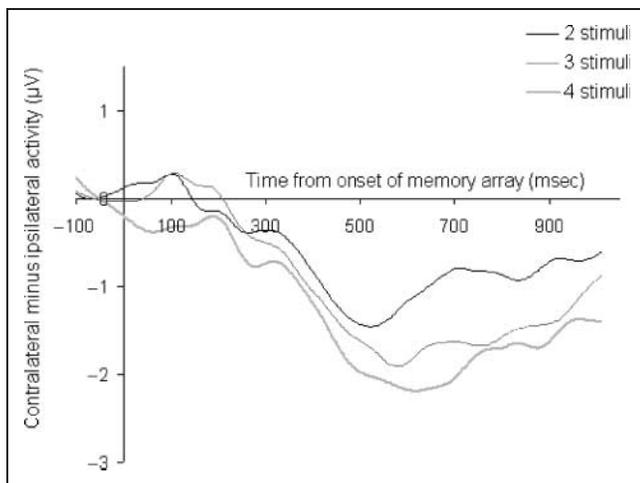
Figure 9. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 3, color condition. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

showed a markedly different pattern, as a function of set size, compared with the color condition. The ANOVA on the SPCN amplitude, with a time window of 450–900 msec after onset of the memory array (with the same variables as in Experiment 1), confirmed that the interaction between Condition and Set Size, $F(2,22) = 4.71$, $p < .05$, $MSE = 0.50$, was significant. The decrease in the SPCN amplitude as the set size increased was significant for colors, $F(1,11) = 15.70$, $p < .005$, $MSE = 0.29$, but for polygons the amplitude with four items was lower relative to two items, a trend that was not significant, $F(1,11) = 1.84$, $p > .20$, $MSE = 0.37$. To test the effect of the pro-
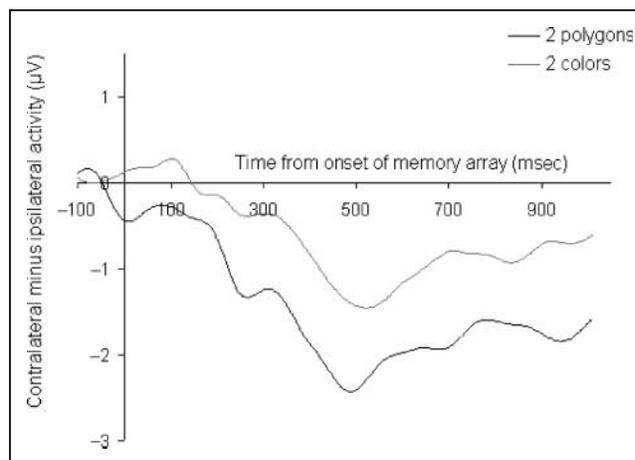
longed memory array interval on the SPCN amplitude, we performed another ANOVA including the results from Experiment 2 and the current one, with Experiment as an independent between-participants variable. Importantly, neither the main effect of Experiment nor any interactions including this variable were close to being significant (all $Fs < 1$). Thus, although tripling the presentation interval did increase mean accuracy, it did not change the SPCN amplitude in any significant way.
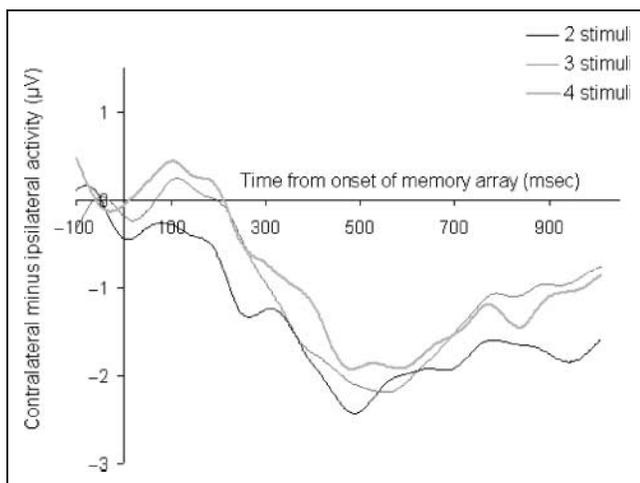
In addition, we analyzed the slopes of the change in amplitude across set size for colors and polygons. The slope for the color condition was -0.44 μV/item, and the slope for the polygons was 0.17 μV/item. The difference between these slopes was significant, $F(1,11) = 11.69$, $p < .01$, $MSE = 0.37$, corroborating our mean amplitude analysis.

As in Experiments 1 and 2, we also compared the SPCN at Set Size 2 for colors and polygons (see Figure 11). The SPCN amplitude was higher for polygons relative to colors, although the number of items was the same, $F(1,11) = 7.95$, $p < .05$, $MSE = 0.57$.

Overall, we replicated the most important patterns of electrophysiological results of Experiments 1 and 2 but with a much longer presentation of the memory array (300 msec rather than 100 msec).

O1/O2. The ANOVA with the same independent variables and with the O1/O2 SPCN amplitude as a dependent variable (see Table 3) yielded a significant main effect of Set Size, $F(2,11) = 3.89$, $p < .05$, $MSE = 0.24$. The interaction between Condition and Set Size was not significant, $F(2,22) = 2.45$, $p = .10$. Inline with the P7/P8 analysis (and previous experiments), the effect of set size was marginally significant for colors, $F(1,11) = 3.96$, $p < .071$, but was far from being significant for polygons, $F < 1$. This time, the difference between two polygons and two colors was significant, $F(1,11) = 6.47$, $p < .05$, $MSE = 0.28$. Thus,

Figure 10. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 3, polygon condition. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

Figure 11. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 3, for two colors and two polygons. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

again the overall numerical pattern was similar to the P7/P8 analysis, although the interaction did not reach a significant level.

Experiment 3 produced results bearing a close resemblance to those of Experiments 1 and 2. For two objects, the SPCN was characterized by a larger amplitude when subjects encoded and maintained more complex objects (polygons) than when they remember simpler objects (colored squares). Moreover, for colors, SPCN increased as the set size increased, but for polygons SPCN was at maximum amplitude already with two objects.

Note that Alvarez and Cavanagh (2004) used an even longer presentation time (of 500 msec), arguing that accuracy reached an asymptotic level only after 450 msec (see also Awh et al., 2007). Thus, it is possible that the 300-msec duration used in Experiment 3 was not enough to completely prevent encoding errors. At this point, we leave it for future research to explore the possibility that with longer exposure durations, the SPCN amplitude for polygons further increases for Set Sizes 3 and 4 (preferably using a within subject design instead of a between experiment comparison).[2]

EXPERIMENT 4

The advantage of using the SPCN as a marker for VSTM capacity in a change detection paradigm is that it constitutes an estimate of VSTM maintenance load uninfluenced by later stages of processing likely involved when comparing memory and test arrays. As argued by Awh et al. (2007), this does not apply to behavioral estimates of performance in the change detection paradigm, which are invariably affected by both limits arising during VSTM maintenance and limits arising during the comparison process. The fact that behavioral and SPCN estimates are influenced by independent stages of processing makes it feasible to hypothesize that a situation in which only the difficulty of the comparison process is selectively manipulated, while the objects' dimension encoded and maintained in VSTM (e.g., colors) is kept constant, should result in a dissociation between SPCN estimates and behavioral estimates.

Experiment 4 was designed to increase the difficulty of the comparison process while keeping constant the load imposed on VSTM. To do so, we compared a low-similarity condition, in which colored squares were displayed using highly distinctive colors (blue, green, and yellow, i.e., the same colors as those used in Experiments 1 and 2) with a high-similarity condition, in which the range of color variation (between blue and green) was much more restricted. When the stimuli are more similar, detecting a change should be more difficult because the magnitude of change is smaller (from one shade of blue to another shade of blue). When the colors are less similar, changes are much larger (e.g., from blue to yellow). Critically, the cause of this difference in difficulty of the overall task should be evident during the comparison processes and

not during retention in VSTM. The reason is that a color constitutes only one feature (whether it is blue or a shade of blue), and thus VSTM should maintain one feature for each color in both conditions. This means that SPCN, as an index of VSTM capacity, should be identical for the high-similarity and the low-similarity conditions, reflecting the number of features that need to be maintained. However, comparing the test and the memory array was expected to be more difficult for similar colors because differences between same and different trials were smaller, and this should be reflected in lower accuracy performance in the high-similarity condition.

A complementary goal of Experiment 4 was to rule out alternative explanations attributing the differences in Experiments 1, 2, and 3 to anticipation for a more difficult test. Perhaps the difference in SPCN amplitude across polygons and colors was due to a difference in the perceived difficulty of the task rather than the complexity of the information that needed to be stored in VSTM. Given that trials in which either shape or color had to be processed were organized in distinct blocks, one could hypothesize that subjects prepared differently and made a greater effort in the more difficult polygon blocks than in the easier color blocks, maybe in anticipation of the more difficult comparison task. This type of alternative explanation of the results of Experiments 1, 2, and 3 would be ruled out if accuracy in the high-similarity condition of Experiment 4 was significantly lower than that in the low-similarity condition, while leaving the SPCN amplitudes unaffected.

Results and Discussion

Behavioral Results

Overall, the accuracy pattern was similar to that found in previous experiments (see Table 1). Accuracy decreased as set size increased, and this effect was more pronounced for similar colors (high similarity) than for distinct colors (low similarity). An ANOVA including the variable Condition (low similarity vs. high similarity) and Set Size (two, four, or six items) on accuracy performance yielded main effects of Condition, $F(1,23) = 468.23$, $p < .0001$, $MSE = 0.002$, and Set Size, $F(2,46) = 383.45$, $p < .0001$, $MSE = 0.001$ (indicating that accuracy dropped by 0.23 from Set Size 2 to Set Size 6). The interaction between Condition and Set Size was also significant, $F(2, 46) = 20.17$, $p < .001$, $MSE = 0.001$, indicating that the drop in accuracy due to set size was more pronounced for low-similarity colors (0.28) than for high-similarity colors (only 0.17). Overall accuracy level was 0.82 for different colors and 0.65 for similar colors.

Although our instructions specifically emphasized accuracy, and mainly to rule out speed–accuracy trade-off, we also analyzed RT (see Table 1). The ANOVA that included the same variables as the accuracy yielded main effects of Condition, $F(1,23) = 24.85$, $p < .005$, $MSE = 18608.13$,
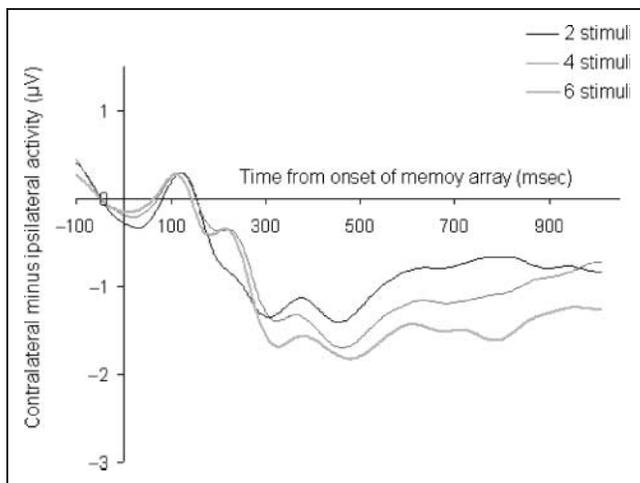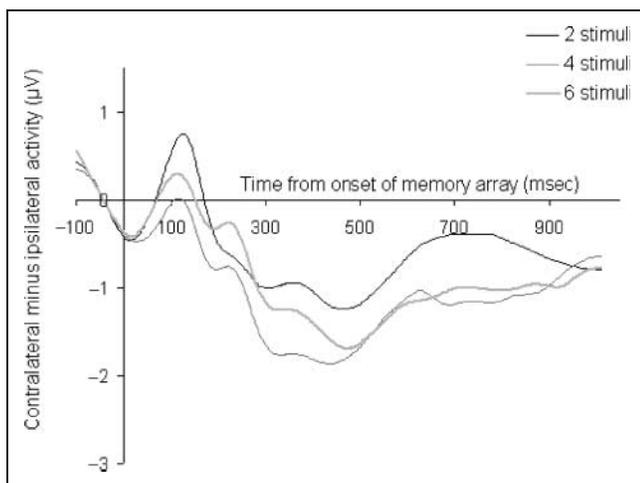
Figure 12. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 4, low-similarity condition. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

and Set Size, $F(2,46) = 16.64$, $p < .001$, $MSE = 19405.04$. The interaction was also significant, $F(2,46) = 5.46$, $p < .05$, $MSE = 3549.89$, indicating that the increase in RT with set size was more pronounced in the low-similarity condition (198 msec), relative to the high-similarity condition (124 msec). The results were not consistent with a speed-accuracy trade-off.

Electrophysiology

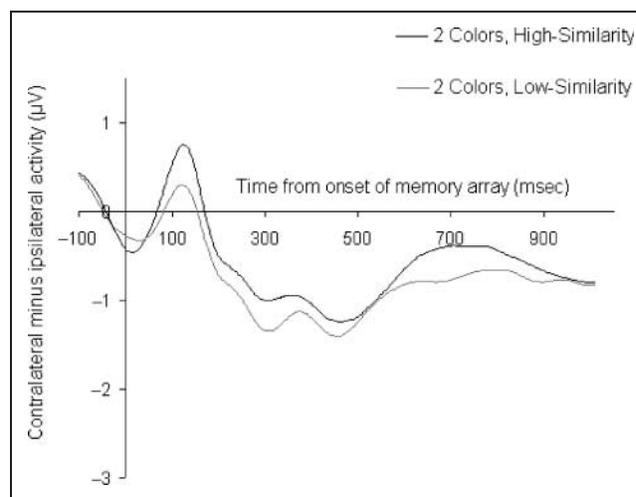*P7/P8.* The grand average subtraction waveforms in the low-similarity condition for each set size are shown in

Figure 12. Figure 13 shows the same waveforms for the high-similarity condition. Table 2 shows the mean SPCN amplitude values for all experimental conditions. We quantified the SPCN by computing the mean amplitude of the contralateral minus ipsilateral difference waveform at electrodes P7/P8 in a time window of 450–900 msec. The mean amplitudes (for each subject for each condition) were submitted to an ANOVA and included the same variables as the accuracy analysis. The only significant effect was that of Set Size, $F(2,46) = 3.35$, $p < .05$, $MSE = 0.45$, which reflected an increase in the amplitude of the SPCN as more colors were to be remembered. Importantly, neither the main effect of Condition, $F(1,23) = 1.28$, $p > .26$, nor the interaction between Condition and Set Size, $F(2,46) = 0.64$, $p > .53$, were significant. Although the interaction was not significant, we directly compared the Load 2 condition across the low-similarity and the high-similarity conditions (as we had done in previous experiments for the colors vs. polygons conditions). The waveforms are displayed in Figure 14. Unlike what we found in Experiments 1, 2, and 3, the amplitude of the SPCN did not differ across the two conditions, $F(1,23) = 0.26$, $p > .61$, suggesting that the storage requirements for low-similarity colors and high-similarity colors was the same. We also analyzed the slope of the change in the amplitude of the SPCN across set size for low-similarity and high-similarity colors. The slope for the low-similarity colors was -0.11 μV/item, and the slope for the high-similarity colors was -0.07 μV/item. The difference between these slopes was not significant, $F < 1$, corroborating our mean amplitude analysis.

As expected, accuracy in the change detection VSTM task was lower in the high-similarity condition than in the

Figure 13. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 4, high-similarity condition. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

Figure 14. Grand average ERP waveforms time locked to the onset of the memory array averaged across P7/P8 electrodes in Experiment 4, for two different colors and two similar colors. For visual purposes, the grand average waveforms were low-pass filtered at 5 Hz (without loss of relevant information, given that the SPCN is a sustained low-frequency wave).

low-similarity condition. This difference in performance could reflect a failure of encoding, retention, and retrieval or in the comparison process between the retrieved memory and the representation of the probe display (Awh et al., 2007). The equivalent mean amplitude of the SPCN waves and the absence of interaction between set size and condition (low vs. high similarity) suggest that, in the present case, performance differences were not in encoding or retention of representations in VSTM. Given the equivalent SPCN results across conditions, the most likely locus (or loci) for the effect of color similarity was that involved in the comparison with the probe display (and associated decision mechanisms).

Importantly, this pattern of results suggests that the SPCN results of Experiments 1, 2, and 3 reflect differences in the storage requirements of polygons and colors rather than differences in overall task difficulty. In Experiment 4, large differences in overall accuracy were found across difficulty levels. Despite these large performance differences, we observed equivalent SPCN amplitudes. These results show that the SPCN was not influenced by the difficulty of the memory comparison task or the anticipated difficulty of the task in general.[3]

O1/O2. The ANOVA with the same independent variables and with the O1/O2 SPCN amplitude as a dependent variable (see Table 3) yielded only a significant main effect of Set Size, $F(2,46) = 7.09$, $p < .05$, $MSE = 0.22$. The interaction between Condition and Set Size was far from being significant, $F < 1$. Inline with the P7/P8 analysis, the effect of Set Size was significant for both similar, $F(1,23) = 5.36$, $p < .05$, $MSE = 0.25$, and different colors, $F(1,23) = 11.62$, $p < .001$, $MSE = 0.15$. The difference between two polygons and two colors was not $F < 1$. Thus, the overall numerical pattern was similar to the P7/P8 analysis.

GENERAL DISCUSSION

The purpose of this study was to measure the capacity of VSTM for simple and complex stimuli using human electrophysiology. In Experiment 1, black polygons served as complex stimuli and colored squares served as simple stimuli. In Experiment 2, only colored polygons were presented, and subjects were encouraged, through explicit instructions, to encode only one feature of the stimuli (either shape or color) in different blocks of trials. Experiment 3 was identical to Experiment 2, but we increased the presentation time of the memory array from 100 to 300 msec. In all experiments, we used the SPCN as an electrophysiological marker of VSTM load during the retention interval. Previous work has shown that the amplitude of the SPCN increases systematically as the amount of information stored in VSTM increases (Jolicœur, Brisson, et al., 2008; Jolicœur, Dell'Acqua, et al., 2007; Perron et al., 2009; McCollough et al., 2007; Vogel et al., 2005; Vogel & Machizawa, 2004) up to the storage capacity of VSTM (Vogel & Machizawa, 2004).

The results of Experiments 1, 2, and 3 were clear-cut: The amplitude of the SPCN increased as more colors were added to the memory array. In contrast, there was no increase in the amplitude of the SPCN as more polygons were to be encoded. This pattern of results suggests that all available capacity in VSTM was required to represent two polygons (producing a saturation of the SPCN at Set Size 2), but that there was additional storage capacity for colors, reflected both by an increase in the amplitude of the SPCN as we increased the number of colors in the memory array and in the estimated capacity of VSTM based on accuracy results. If a larger SPCN indicates that more storage capacity has been engaged, in a given condition or by a given type of stimulus, then our results suggest that more capacity was used up when participants encoded and maintained the shape of a polygon than when they encoded and maintained the color of an object. It is possible that neurons mediating VSTM need to work harder to maintain more complex objects, which in turn is consistent with the view that VSTM capacity depends also on stimulus complexity (Alvarez & Cavanagh, 2004). In addition, Experiment 4 demonstrated that the larger SPCN amplitude found in Experiments 1, 2, and 3 was not a result of preparation toward a difficult retrieval/memory comparison task.

Note that, by monitoring the SPCN, we avoided the criticism that was raised by Awh et al. (2007) who attributed the low capacity for complex objects to an error-prone comparison process between the memory and the test arrays. The reason is that we measured the SPCN during the retention interval, before the presentation of the test array, and thus before the comparison process. The present results, however, are in perfect accord with and support the hypothesis of Awh et al. that the type of test/memory array comparison process plays a role under experimental conditions analogous to the present ones. In fact, Experiment 4 showed that the comparison process is presumably responsible for the poor accuracy performance when the memory and test arrays were very similar, although VSTM capacity was the same for both color conditions. It is thus likely that the comparison process was part of the cause for the poor accuracy in the random polygon conditions, although more capacity was demanded for the maintenance of more complex information. In any case, our results support the notion that VSTM storage capacity, per se, is sensitive to stimulus complexity, based on direct electrophysiological measures of VSTM retention activity (ruling out error-prone comparison process as a possible explanation).

Woodman and Vogel (2008) found that when subjects remember objects with important contour information (orientations in that case), this resulted in a significant increase in overall amplitude of the SPCN, irrespective of set size, relative to an equivalent number of colors. Given that orientation was not particularly complex in their work (showing identical behavioral performance with color) and did not interact with set size, they concluded

that this increase in amplitude was likely due to some form of sensitivity to the nature of the stimulus rather than to storage capacity differences. On the basis of this work, one interpretation of the greater SPCN amplitude found for two polygons than for two colors (see Figures 5, 8, and 11) may be that orientation or shape features simply produce a greater electrophysiological response than color. However, the present results go beyond these prior demonstrations in a critical way. Whereas the SPCN amplitude increased consistently with increasing memory set size in the case of color, the SPCN amplitude either leveled off or tended to decrease with increasing memory set size in the case of polygons. This was consistently reflected in significant interactions between set size and stimulus complexity found across Experiments 1–3 when assessed directly on mean SPCN amplitudes and also on a quantification of the slopes of the set size effects across different levels of stimulus complexity.

But why would a polygon consume more storage capacity in VSTM than a color? Presumably, polygons are made up of a collection of shape features, which would imply that remembering a polygon involves the encoding and the retention of several within-dimension feature conjunctions. Memory for within-dimension conjunctions has been shown to be a particularly demanding process (Wheeler & Treisman, 2002). This may be different when encoding colored squares, which can be represented as unique features, or colored, oriented bars that required the encoding and conjunction of two features across different dimensions.

Interestingly, in a recent fMRI study, Xu and Chun (2006) have revealed two dissociable neural mechanisms mediating VSTM in the human brain. They hypothesized that the inferior intraparietal sulcus is specialized to represent a fixed number of objects, regardless of complexity, whereas neurons in the superior intraparietal sulcus and in the lateral occipital complex were hypothesized to be specialized to represent the total amount of the visual information encoded, being therefore sensitive to object complexity. In this perspective, the present results are broadly consistent with the second system postulated by Xu and Chun, insofar as the SPCN amplitude was sensitive the stimulus complexity.

One speculation that goes beyond what our results firmly support is that there may be an efficient trade-off between storage resolution and capacity. When only one object is presented, resolution can be high even for a complex stimulus composed of multiple within-dimension feature conjunctions, leaving performance at near perfect levels (Awh et al., 2007; Alvarez & Cavanagh, 2004). An additional cost in terms of storage capacity may compensate for resolution (at least when one object is presented). When up to three or four simple objects are maintained, resolution may remain a nonlimiting factor, and performance is determined mainly by the number of objects. However, when several complex stimuli are presented, resolution may become a serious limiting factor because

each high-resolution item requires high-storage capacity. Each high-resolution representation would consume more storage capacity relative to the capacity for simple object, and this would reduce the total number of objects that can be maintained in VSTM. This means that the amplitude of the SPCN indicates that more effort is devoted to maintain objects in VSTM and that this activity could reflect the number of features (for single-feature objects) or perhaps the number of conjunctions of features (objects) when the features belong to different dimensions. However, for objects composed of multiple within-dimension feature conjunctions (such as the polygon shapes we used in the present work, shown in Figure 1B), the amplitude of the SPCN would reflect the overall load on the VSTM system, with load determined by the number and resolution of the represented objects.

A possible objection to the explanation we are proposing relates to the structure of the paradigms used in the present study. Specifically, given that complexity conditions were invariably blocked in the designs of Experiments 1–4, one may wonder whether this feature could have artificially induced the observed differences in memory set size effects on the SPCN across levels of complexity. In other words, given that subjects knew what type of stimuli was about to be presented, it is possible that they may have strategically chosen to encode only a subset of the memory arrays in the case of difficult-to-encode stimuli (i.e., polygons, only two of them in all cases), but they may have attempted to encode all of them when the task required encoding and remembering easy-to-encode stimuli (i.e., colors). An obvious counterpoint to this view, however, is provided by the results of Experiment 4, where sets of colors characterized by either low similarity or high similarity were also presented in different blocks. That the two conditions differed in difficulty was corroborated by the large difference in accuracy across difficulty levels in Experiment 4, with a markedly worse performance in the change detection task in the high-similarity condition relative to the low-similarity condition. If blocking task difficulty invariably induced subjects to encode only a subset of the items in the more difficult condition, then SPCN amplitude should have had a flat memory set size function in the high-similarity condition of Experiment 4, much in the same way as in the polygon conditions of Experiments 1–3. This was not the pattern found in Experiment 4. A sizable increase in SPCN amplitude was found for both levels of the difficulty manipulation, together with the absence of an interaction between set size and difficulty.

Finally, the Experiment 4 also suggests that one cannot explain the results of Experiments 1, 2, and 3 based on task difficulty (as assessed by behavioral performance). As can be seen in Table 1, task difficulty as assessed by accuracy or RT was equivalent, or greater, in the high-similarity condition of Experiment 4 as in the polygon conditions of Experiments 1–3. It is for these considerations that the fact that there was little or no increase in

SPCN amplitude for arrays of three or four polygons is particularly interesting and strongly support our claim that the most parsimonious account of the results is that polygons (which are composed of multiple within-dimension shape conjunctions) tend to saturate VSTM capacity even when only two polygons must be retained in memory. This, in turn, explains why SPCN amplitude could not increase as additional polygons were added in the memory array, even when, as we did in Experiment 3, we minimized the possibility that processing of the memory array was bottlenecked at encoding by increasing the exposure duration of the memory array from 100 msec (Experiments 1 and 2) to 300 msec. Despite a very difficult memory task in the high-similarity condition of Experiment 4, each color took up no more or less storage capacity as any other, and so VSTM storage capacity did not saturate at Set Size 2, as it did for polygons.

APPENDIX

In each experiment, we estimated the number of items available in VSTM using a formula developed by Cowan (2001) and Pashler (1988). These investigators developed a simple equation for estimating the number of items, K, that are available in working memory in a change detection task: $K = S \times (\text{Hit} + \text{CR} - 1)$, when CR is the correct rejection rate, Hit is the hit rate, and S is the number of items composing the array.

In Experiment 1, for the group without the silent rehearsal manipulation, mean K for colors was 1.88, 2.60, and 3.00 items when two, three, and four stimuli were presented, respectively, and 1.01, 0.83, and 0.90 items when two, three, and four random polygons were presented, respectively. In the silent rehearsal group, K for color was 1.91, 2.64, and 3.06 items when two, three, and four stimuli were presented, respectively, and 1.01, 0.96, and 0.86 items when two, three, and four random polygons were presented, respectively.

In Experiment 2, in the color condition, K was 1.86, 2.63, and 3.05 when two, three, and four colors were presented, respectively. In the polygon condition, K was 0.89, 0.86, and 0.85 items when two, three, and four stimuli were presented, respectively.

In Experiment 3, in the color condition, K was 1.75, 2.29, and 2.64 when two, three, and four colors were presented, respectively. In the polygon condition, K was 1.02, 0.90, and 0.88 items when two, three, and four stimuli were presented, respectively.

In Experiment 4, K was 1.81, 2.75, and 2.77 in the low-similarity condition, and 1.28, 1.06, and 0.92 in the high-similarity condition for Set Sizes 2, 4, and 6, respectively.

Notes

1. When eliminating two participants (with a slope of 3.94 and 1.84), the average slope of all other participants was 0.19.
2. In our own pilot work using 500 msec duration, many subjects found it difficult to resist the urge to move their eyes toward the target shapes, resulting in too many discarded trials due to eye movements.
3. We assume there is no interference/competition between stimuli during the VSTM retention interval. Nonetheless, if such competition did occur, it would be expected to be higher when the stimuli are more similar. Thus, one would expect larger SPCN effects in the high-similarity condition. However, we did not observe such an effect.

REFERENCES

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short term memory is set both by visual information load and by number of objects. *Psychological Science, 15,* 106–111.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items, regardless of complexity. *Psychological Science, 18,* 622–628.

Brisson, B., & Jolicœur, P. (2007a). A psychological refractory period in access to visual short-term memory and the deployment of visual-spatial attention: Multitasking processing deficits revealed by event-related potentials. *Psychophysiology, 44,* 323–333.

Brisson, B., & Jolicœur, P. (2007b). Electrophysiological evidence of central interference on the control of visual-spatial attention. *Psychonomic Bulletin & Review, 14,* 126–132.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185.

Dell'Acqua, R., Sessa, P., Jolicœur, P., & Robitaille, N. (2006). Spatial attention freezes during the attentional blink. *Psychophysiology, 43,* 394–400.

Delvenne, J. F., & Bruyer, R. (2004). Does visual short-term memory store bound features? *Visual Cognition, 11,* 1–27.

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review, 12,* 1127–1133.

Jolicœur, P., Brisson, B., & Robitaille, N. (2008). Dissociation of the N2pc and sustained posterior contralateral negativity in a choice response task. *Brain Research, 1215C,* 160–172.

Jolicœur, P., Dell'Acqua, R., Brisson, B., Robitaille, N., Sauvé, K., Leblanc, E., et al. (2007). Visual spatial attention and visual short-term memory: Electro-magnetic explorations of the mind. In V. Coltheart (Ed.), *Tutorials in visual cognition.* Hove, UK: Psychology Press.

Jolicœur, P., Sessa, P., Dell'Acqua, R., & Robitaille, N. (2006). On the control of visual spatial attention: Evidence from human electrophysiology. *Psychological Research, 70,* 414–424.

Klaver, P., Talsma, D., Wijers, A. A., Heinze, H.-J., & Mulder, G. (1999). An event-related brain potential correlate of visual short-term memory. *NeuroReport, 10,* 2001–2005.

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

McCollough, A. W., Machizawa, M. G., & Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex, 43,* 77–94.

Olson, I., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong object" hypothesis. *Perception & Psychophysics, 64,* 1055–1067.

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 8776–8780.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics, 44,* 369–378.

Perron, R., Lefebvre, C., Robitaille, N., Brisson, B., Gosselin, F., Arguin, M., et al. (2009). Attentional and anatomical considerations for the representation of simple stimuli in visual short-term memory: Evidence from human electrophysiology. *Psychological Research, 73,* 222–232.

Pivik, R. T., Broughton, R. J., Coppola, R., Davidson, R. J., Fow, N., & Nuwer, M. R. (1993). Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts. *Psychophysiology, 30,* 547–558.

Robitaille, N., & Jolicœur, P. (2006). Fundamental properties of the N2pc as an index of spatial attention: Effects of masking. *Canadian Journal of Experimental Psychology, 60,* 79–89.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74* (11, Whole No. 498).

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751–753.

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

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

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 92–114.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131,* 48–64.

Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin & Review, 15,* 223–229.

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