Obligatory encoding of task-irrelevant features depletes working memory resources

Louise Marshall
Sobell Department of Motor Neuroscience and Movement Disorders, UCL Institute of Neurology, London, UK

Paul M Bays
Sobell Department of Motor Neuroscience and Movement Disorders, UCL Institute of Neurology, London, UK

Selective attention is often considered the "gateway" to visual working memory (VWM). However, the extent to which we can voluntarily control which of an object's features enter memory remains subject to debate. Recent research has converged on the concept of VWM as a limited commodity distributed between elements of a visual scene. Consequently, as memory load increases, the fidelity with which each visual feature is stored decreases. Here we used changes in recall precision to probe whether task-irrelevant features were encoded into VWM when individuals were asked to store specific feature dimensions. Recall precision for both color and orientation was significantly enhanced when task-irrelevant features were removed, but knowledge of which features would be probed provided no advantage over having to memorize both features of all items. Next, we assessed the effect an interpolated orientation-or color-matching task had on the resolution with which orientations in a memory array were stored. We found that the presence of orientation information in the second array disrupted memory of the first array. The cost to recall precision was identical whether the interfering features had to be remembered, attended to, or could be ignored. Therefore, it appears that storing, or merely attending to, one feature of an object is sufficient to promote automatic encoding of all its features, depleting VWM resources. However, the precision cost was abolished when the match task preceded the memory array. So, while encoding is automatic, maintenance is voluntary, allowing resources to be reallocated to store new visual information.

Introduction

As the number of items in a visual scene increases, the fidelity with which each item is stored in visual working memory (VWM) declines (Alvarez & Cavanagh, 2004; Bays & Husain, 2008; Lakha & Wright, 2004; Palmer, 1990; Wilken & Ma, 2004). This observation has been interpreted as reflecting the limited nature of VWM resources, which must be distributed between elements of a visual scene. The resource may be continuous (Bays, Catalao, & Husain, 2009; Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011a; Bays & Husain, 2008; Elmore et al., 2011; Gorgoraptis, Catalao, Bays, & Husain, 2011; Huang, 2010; Lara and Wallis, 2012; Van den Berg, Shin, Chou, George, & Ma, 2012) or divided into a small number of discrete chunks or quanta (Anderson, Vogel, & Awh, 2011; Zhang & Luck, 2008, 2009). In either case, a larger memory load means that VWM resources must be distributed among a greater number of items, reducing the precision with which any individual item can be recalled.

Because every additional item stored has a cost in terms of recall fidelity, this finding places renewed emphasis on the ecological importance of controlling which elements of the visual environment gain access to memory. Selective attention is often conceptualized as the "gateway" to VWM, controlling which items in a visual scene VWM resources are allocated to. However, the extent to which individuals can voluntarily control which of an object's particular features enter memory remains less clear.

Previously, the dominant view of VWM was of a small number of independent memory "slots" (typically four), each storing visual information related to a distinct object (Cowan, 2001; Luck & Vogel, 1997; Pashler, 1988; Sperling, 1960; Vogel, Woodman, & Luck, 2001). Within this framework of a fixed capacity limit, there has been considerable debate as to whether the encoding of visual information into memory occurs at an object- or feature-level. Luck and Vogel (1997) showed that increasing the number of objects in a memory array increased recall errors, but increasing the number of features per object did not. This finding was interpreted as evidence for "whole object" representation in VWM, mirroring results in the attention

Citation: Marshall, L., & Bays, P. M. (2013). Obligatory encoding of task-irrelevant features depletes working memory resources. *Journal of Vision, 13*(2):21, 1-13, http://www.journalofvision.org/content/13/2/21, doi:10.1167/13.2.21.

doi: 10.1167/13.2.21 Received September 20, 2012; published February 18, 2013 ISSN 1534-7362

literature that suggested objects were the units of attentional selection (Duncan, 1984; O'Craven, Downing, & Kanwisher, 1999; Vecera & Farah, 1994).

However, subsequent research (Olson & Jiang, 2002; Wheeler & Treisman, 2002; Xu, 2002) failed to support a memory advantage of grouping features into objects, instead finding that error rates were determined by the total number of features that needed to be remembered within each feature dimension (e.g., the total number of colors in a memory array, regardless of whether they belonged to one or several objects). Instead of a single memory store maintaining integrated object representations, Wheeler and Treisman (2002) proposed parallel memory stores for each feature dimension, with independent capacities. This account finds support in behavioral and neurophysiological studies showing that attention can be feature- as well as object-based (Martinez-Trujillo & Treue, 2004; Maunsell & Treue, 2006; Saenz, Buracas, & Boynton, 2003).

In recent years, the focus of VWM research has shifted from binary (correct/incorrect) measures of memory performance to instead examine the way recall errors are distributed in the space of possible responses. The observation that the variability of recall error increases monotonically with the number of competing items in memory (Bays & Husain, 2008; Lakha & Wright, 2004; Palmer, 1990; Wilken & Ma, 2004; Zhang & Luck, 2008) has proven difficult to reconcile with models of VWM in which each object is stored in a separate "slot," instead suggesting that VWM is a limited commodity distributed between objects.

Furthermore, some initial results using this newer methodology seem easier to accommodate within a feature- rather than object-based account of VWM. When subjects were required to reproduce from memory two features of one object in a memory array (e.g., color and orientation), recall errors were found to occur independently in each feature dimension (Bays, Wu, & Husain, 2011b; Fougnie & Alvarez, 2011). This result is incompatible with the claim that errors arise from a failure to store or maintain whole objects (Cowan, 2001; Luck & Vogel, 1997; Zhang & Luck, 2008, 2009). Instead this dissociation could imply that VWM resources are independently allocated to elements of a visual scene at the feature level.

Here we use changes in recall precision as a probe to investigate the conditions under which an object's features enter memory. Since recall precision declines as memory load increases (Bays & Husain, 2008; Lakha & Wright, 2004; Palmer, 1990; Wilken & Ma, 2004; Zhang & Luck, 2008), an increase in the precision with which a visual feature (e.g., an object's color) is recalled provides evidence for a decrease in the number of competing features (e.g., colors of other objects) in memory. Using this approach, we first examine whether participants can voluntarily store one feature of an object in memory without allocating VWM resources to its other features. We then investigate the consequences of attending to features of an object for which there is no memory requirement. Our results indicate that encoding of task-irrelevant features into VWM is largely obligatory, whereas subsequent maintenance of visual information is under voluntary control.

General methods

Procedure

A total of 32 subjects (16 male, 16 female; aged 18–36 years) participated in the study after giving informed consent, in accordance with the Declaration of Helsinki. All subjects reported normal color vision and had normal or corrected-to-normal visual acuity. Stimuli were presented on a 21" CRT monitor with a refresh rate $\geq$ 130 Hz. Subjects sat with their head supported by a forehead- and chin-rest and viewed the monitor at a distance of 60 cm. Eye position was monitored online at 1000 Hz using an infrared eye tracker (SR Research, Ontario, Canada).

In all experiments, a trial began with the presentation of a central white fixation cross (0.75° of visual angle) against a black background. Once a stable fixation was recorded within 2° of the cross, a sequence of two displays, each followed by a blank interval, was presented. Each display was presented for 1000 ms and consisted of either two colored circles (diameter 4°) or two oriented bars (0.75° $\times$ 4°) positioned on an imaginary circle with an 8° radius around the point of fixation (examples in Figure 1). Positions of the four items presented across the two displays were assigned by selecting a random location on the circle for the first item, then randomly allocating the locations 90°, 180°, and 270° from this item to the remaining three items. This procedure ensured that the four items were equally distributed around fixation but their locations were otherwise random. Each item's color was independently chosen at random from a color wheel (defined by a circle in CIE $L^*a^*b^*$ space with center at $a^* = b^* = 20$, radius 60, and constant luminance $L^* = 50$). Bar orientations were independently chosen at random from the full range of possible orientations (0°–180°).

At the end of each trial, subjects were presented with a single *probe* item positioned at a location previously occupied by an item in one of the preceding memory arrays. The presentation of a white probe bar indicated that an orientation recall was required, while a colored probe circle instructed the subject to recall color. Subjects used an input dial (PowerMate USB Multimedia Controller, Griffin Technology, USA) to adjust the orientation or color of the probe item to match the

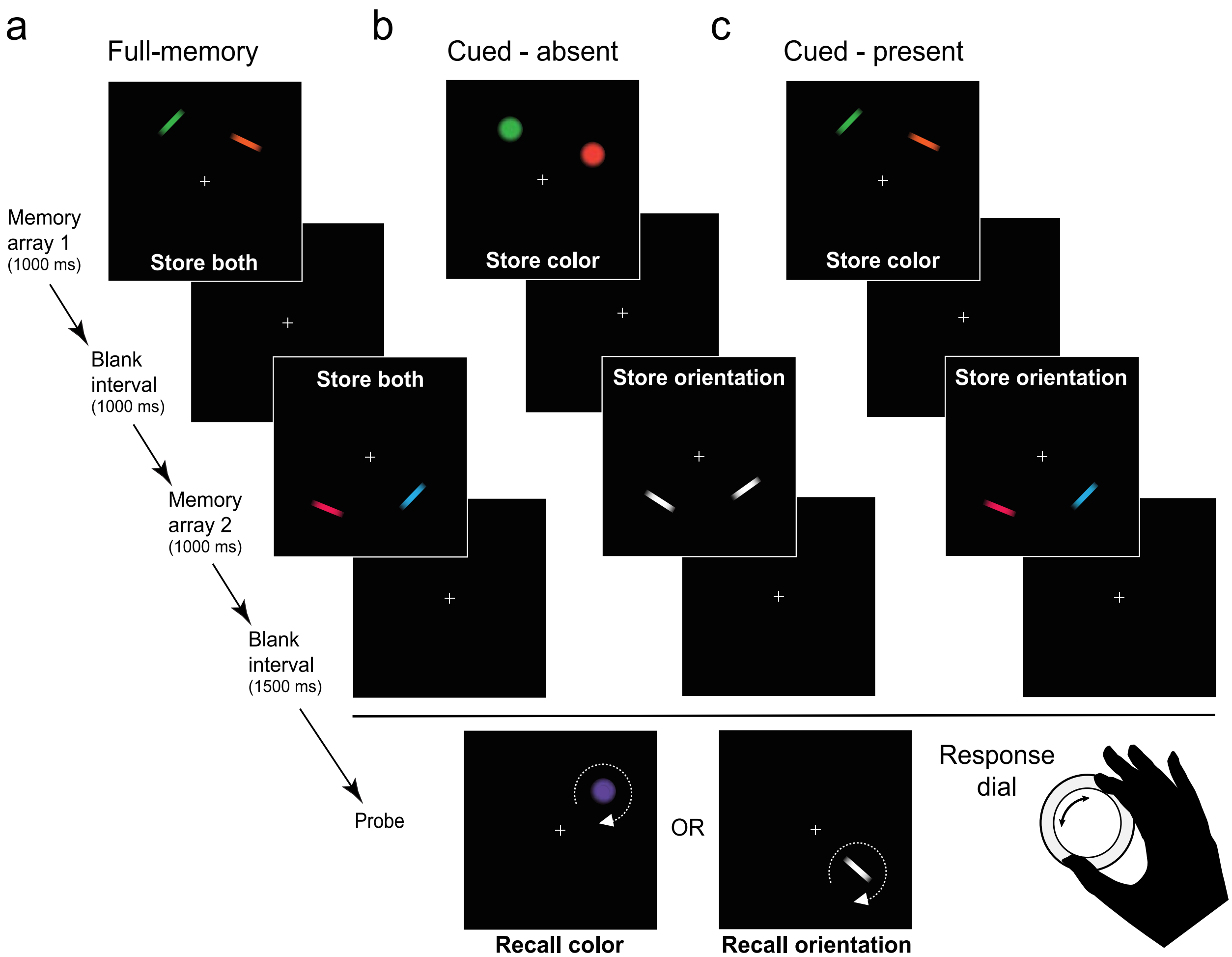

Figure 1. The recall task used in Experiment 1. (a) In the *full-memory* condition, subjects were presented with two sequential arrays of colored, oriented bars and instructed to remember all of the colors and orientations they saw. The four items presented across the two arrays were randomly allocated to four positions, each separated by 90°, on an imaginary circle centered on fixation. After a blank interval, a probe item was presented: either a colored circle, indicating that color memory would be tested, or a white bar, indicating memory for orientation would be tested. In each case, subjects used a response dial to adjust the color/orientation of the probe to match the color/orientation of the bar that had appeared in the same location in the preceding displays. (b) In the *cued-absent* condition, the memory load was halved and task-irrelevant information removed from the displays: one array contained only color information (colored circles) and the other array contained only orientation information (white bars). (c) The *cued-present* condition was visually-identical to the *full-memory* condition but, like in the *cued-absent* condition, subjects had to remember only the colors of one array and only the orientations of the other array.

remembered orientation or color of the item that had been presented at the same location (the *target*). The probe's initial orientation or color was randomly assigned. Responses were not timed and subjects were instructed to be as precise as possible. Any trial on which gaze deviated more than 2° from the central cross during the sequence of presentations preceding the recall display was aborted and restarted with new feature values.

Analysis

A measure of recall error was obtained on each trial by calculating the angular deviation between the target item's orientation and the orientation reported by the subject, or the angular deviation between the target item's color on the color wheel and the color actually reported. Color and orientation responses were analyzed in terms of the circular parameter space of possible feature values (ranging from $-\pi$ to π radians). For each combination of subject and experimental condition, we calculated a measure of recall *precision*, defined here as the reciprocal of the standard deviation of the error. As in previous studies (Bays et al., 2009; Bays et al., 2011a; Bays et al., 2011b; Gorgoraptis et al., 2011), we used Fisher's definition of standard deviation for circular data (Fisher, 1995) and subtracted from the precision estimate the value expected by chance (i.e., if the subject had responded at random on each trial). Statistical comparisons were made between experimental conditions using repeated-measures analysis of variance (ANOVA) and paired-samples t-tests.

Two models were considered to describe the distribution of responses relative to the target feature value: (1) a mixture of a Von Mises and uniform distribution (parameterized by σ, the circular standard deviation of a Von Mises distribution centered on the target value, and P_m, the mixing proportion; Zhang & Luck, 2008), and (2) a wrapped stable distribution centered on the target value with zero skew (parameterized by α, which determines the kurtosis, and γ which determines the standard deviation). The latter represents a generalization of the Gaussian distribution on the circle to symmetric distributions of variable kurtosis, with $\alpha = 2$ corresponding exactly to the wrapped Gaussian distribution (Arthur, 2008). Maximum likelihood parameters were obtained for each model based on the Nelder-Mead method (*fminsearch* in Matlab), and models compared on the basis of the Bayesian Information Criterion (BIC).

To provide non-parametric tests for the influence of non-targets on recall errors (Bays et al., 2009; Bays et al., 2011b), we examined the deviation of responses from each non-target feature value on each trial. Central tendency in this distribution was tested using the V test for circular data (Zar, 2010), where a positive result indicates that responses were significantly clustered around non-target feature values. Effects of experimental condition on this distribution were tested by two-sample Kolmogorov-Smirnov tests.

Experiment 1

Our first experiment aimed to test whether participants could voluntarily control which features of a set of visual objects entered working memory.

Methods

12 subjects (five male, seven female; aged 21–36 years) participated in Experiment 1. The precision with which participants could recall from memory the orientation and/or color of different items was assessed in three different conditions. Subjects were tested on each condition in a separate block of 80 trials, and the order in which the three conditions were performed was counterbalanced across subjects. The procedure is illustrated in Figure 1.

In the *full-memory* condition (Figure 1a), subjects were presented with two sequential pairs of randomly oriented and colored-bars with the instruction to memorize both features of all four items. On each trial, they could be asked to recall either the orientation or the color of any one of the four items. There were, therefore, eight task-relevant features to memorize on each trial. The item to be recalled (the target) was selected at random, such that on 50% of trials the target item was from the first memory array and on 50% it was from the second array. Subjects were asked to report the orientation on half of the trials and color on the other half.

In the *cued-absent* condition (Figure 1b), the number of visual features displayed across the two memory arrays was halved. On each trial, subjects were presented with two randomly colored circles in one array and two randomly oriented white bars in the other, making four task-relevant features in total. The target item was selected at random, such that on 50% of trials subjects were required to recall the orientation of one of the bars, and on the remaining 50% they had to recall the color of one of the circles. For each subject, the order in which the color and orientation memory arrays were presented was the same for all trials, but order was counterbalanced across subjects.

In the *cued-present* condition (Figure 1c), the two memory arrays were visually identical to those presented in the *full-memory* condition, but the task-relevant memory load was the same as that in the *cued-absent* condition. Subjects were instructed to remember only the orientations of bars presented in one display (e.g., the first) and only the colors of bars presented in the other display (e.g., the second), and to ignore the other features. This order was kept the same throughout the block of trials (but counterbalanced across subjects). Probe selection was consistent with these instructions, so although eight features were presented, subjects only had to memorize four features on each trial to successfully complete the task (two orientations from one array and two colors from the other). The additional feature information present was task-irrelevant.

Before starting each new block of trials, subjects completed 10 practice trials to familiarize them with the displays and instructions specific to the condition. These practice trials were discarded from analysis.

We used sequentially presented arrays with a relatively long exposure duration in order to ensure that the different items competed for memory rather than encoding resources (Bays et al., 2009; Bays et al., 2011a), and to provide participants with a salient cue as to which features to remember. We varied the number of to-be-remembered features *within* each feature dimension (e.g., four colors vs. two colors) rather than varying the number of feature dimensions that had to be remembered (e.g., color and orientation vs. color only) because of evidence from the change detection literature that features from different dimensions do not compete for storage in working memory (Luck &

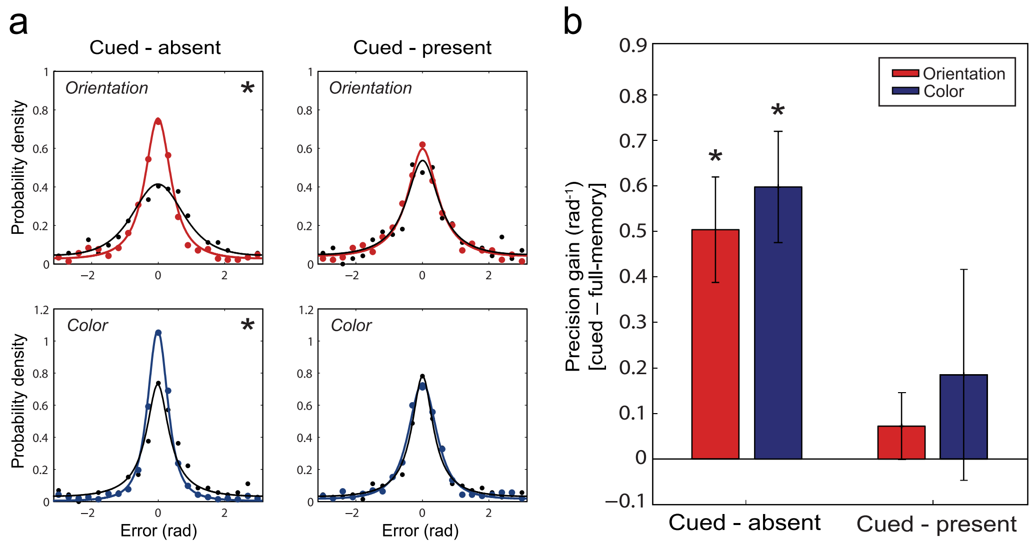

Figure 2. Error distributions and recall precision in Experiment 1. (a) Distribution of errors in the *cued-absent* (left) and *cued-present* (right) conditions for recall of orientation (red) and color (blue). Error distributions for the corresponding trials in the *full-memory* condition are shown in black. Data points indicate mean probability over subjects. Curves correspond to subject-averaged best-fitting wrapped stable distributions (see Methods). Asterisks indicate conditions where precision was significantly enhanced relative to the *full-memory* condition ($p < 0.05$). (b) Mean difference in recall precision (*precision gain*) between each of the *cued* conditions and the *full-memory* condition. Error bars indicate ±1 *SE*. Asterisks indicate significant precision gains ($p < 0.05$).

Vogel, 1997; Olson & Jiang, 2002; Vogel et al., 2001; Wheeler & Treisman, 2002; but see Fougnie, Asplund, & Marois, 2010).

Analysis

Recall precision was calculated separately for each subject, experimental condition, and feature dimension (orientation or color). To assess whether subjects were successful in memorizing task-relevant features of an object without allocating memory resources to its other features, we compared precision in the *cued-absent* and *cued-present* conditions (when subjects only had to memorize one feature dimension in each display) to precision in the *full-memory* condition (when the feature memory load was doubled). Specifically, for each subject, we subtracted from the precision measures obtained in *cued-absent* and *cued-present* conditions the precision calculated for the same feature (color or orientation) presented in the equivalent display (first or second memory array) in the *full-memory* condition. This calculation provided a measure of the *precision gain* associated with reducing the number of task-relevant features in each of the *cued* conditions.

Results and discussion

Figure 2a plots the distribution of orientation (red) and color (blue) errors in each of the *cued* conditions, along with the errors made in the corresponding *full-memory* conditions (black). In the *cued-absent* condition, where the feature memory load was halved and task-irrelevant information was absent from the arrays, there was a decrease in variability of the response errors compared to the *full-memory* condition, as indicated by the taller, narrower response distributions (red/blue versus black in Figure 2a, left). A direct measure of the *precision gain*, the difference in recall precision between *cued-absent* and *full-memory* conditions, is plotted for each feature dimension in Figure 2b (left). The enhanced precision in the *cued-absent* condition, in which fewer features were presented, is consistent with the well-documented relationship between increasing memory load and decreasing precision of recall (Bays & Husain, 2008; Palmer, 1990; Wilken & Ma, 2004; Zhang & Luck, 2008).

By contrast, response distributions in the *cued-present* condition, where feature load was halved but task-irrelevant information was present in the arrays, closely resembled responses in the *full-memory* condition (red/blue versus black in Figure 2a, right), and there was no consistent precision gain in this condition (Figure 2b, right). This finding suggests that the presence of additional, task-irrelevant feature information in the memory arrays of the *cued-present* task caused an increase in the variability of errors comparable to that observed in the *full-memory* task when twice the number of features had to be stored.

Statistical analysis of the precision gain for each of the *cued* conditions relative to the *full-memory* condition confirmed this pattern of results. A repeated measures ANOVA, with *condition* (*cued-absent* or *cued-present*) and *feature* (color or orientation) as within-subjects factors, revealed no significant effect of feature, $F(1, 11) = 0.002$, $p = 0.97$, but a significant effect of condition, $F(1, 11) = 8.8$, $p = 0.013$. Recall precision for both feature dimensions was significantly enhanced in the *cued-absent* condition compared to the *full-memory* condition, color: $t(11) = 4.8$, $p = 0.001$; orientation: $t(11) = 4.7$, $p = 0.001$. No significant precision gain was observed in the *cued-present* condition, color: $t(11) = 0.17$, $p = 0.87$; orientation: $t(11) = 1.9$, $p = 0.078$.

Therefore, recall precision for both color and orientation was significantly enhanced, relative to the *full-memory* condition, only when task-irrelevant features were removed from the memory arrays, as in the *cued-absent* condition. The presence of additional features in the *cued-present* condition had an equivalent deleterious effect on recall precision as when they had to be stored in memory, despite explicit instructions to ignore them. It seems, therefore, that subjects were unable to take advantage of the knowledge of which features of an object were relevant to the memory task, despite the cost in task performance that resulted from it. These results suggest that, when an individual attempts to store specific features of a visual object in memory, there is concurrent obligatory encoding of other features belonging to that object. These task-irrelevant features deplete VWM resources to the same extent as explicitly memorized features, with a corresponding cost to the precision with which features from the same dimension belonging to other objects can be recalled.

Consistent with the results of previous studies that have investigated reproduction or report of visual features from memory (Anderson et al., 2011; Bays et al., 2009; Bays & Husain, 2008; Van den Berg et al., 2012; Zhang & Luck, 2008), we found that the distributions of errors in our data (Figure 2a) did not precisely follow a Gaussian distribution (or its circular equivalent, the Von Mises distribution). While the exact distribution of responses is not vital for interpreting the main results of this experiment, which depend only on the observation that variability increases with memory load, we consider it may be of some interest to examine these distributions in more detail.

Several hypotheses have been put forward to explain the divergence of errors from a Von Mises distribution. One proposal is that the error distribution consists of a mixture of a uniform and a Von Mises distribution (Anderson et al., 2011; Zhang & Luck, 2008); this suggestion is related to the hypothesis that working memory resources could be *quantized*, i.e., divided into a small number of discrete chunks that are distributed between objects. A second possibility is that errors are distributed according to a continuous distribution with higher kurtosis (i.e., heavier tails) than the Von Mises. One way in which such a distribution could arise is from an infinite mixture of Von Mises distributions of different widths (Van den Berg et al., 2012). This latter suggestion is linked to the proposal that working memory resources are continuous, but there is variability in their allocation (i.e., resources are not necessarily evenly distributed between objects).

For the present data, a wrapped stable distribution (a generalization of the circular Gaussian distribution with variable kurtosis) was found to provide a marginally better fit overall than a Von Mises-uniform mixture (relative BIC, 2.5). Curves plotted in Figure 2a correspond to the best-fitting wrapped stable distributions.

A further important consideration is the presence of *non-target* responses. These are instances where the subject accurately reproduces the color or orientation of the wrong item, i.e., one of the items presented on the trial other than the target item. These responses may arise as a result of variability in memory for the probe feature (here location), or as a result of errors in maintaining the binding information that links features of an object together (Bays et al., 2009; Bays et al., 2011b). While these responses appear uniformly-distributed relative to the target feature, their presence can be detected on the basis of a clustering of responses around non-target feature values. Consistent with this situation, we found significant central tendency in the deviation of responses from non-target feature values (V test, $p < 0.001$), indicating that non-target responses did contribute to errors in our task. However, we found no significant differences in this non-target error distribution between the different conditions of our experiment (Kolmogorov-Smirnov test, $p > 0.11$), indicating that changes in the frequency of these responses are unlikely to have contributed substantially to our main findings.

Experiment 2

The results of Experiment 1 provide evidence for involuntary storage of task-irrelevant features when subjects were required to memorize just one of the features of a visual object. In Experiments 2A and 2B we examined whether merely attending to an item or feature, with no requirement to remember it, is sufficient to promote storage of task-irrelevant features in VWM.

Methods

12 subjects (six male, six female; aged 19-36 years; all right-handed) participated in Experiment 2A. We

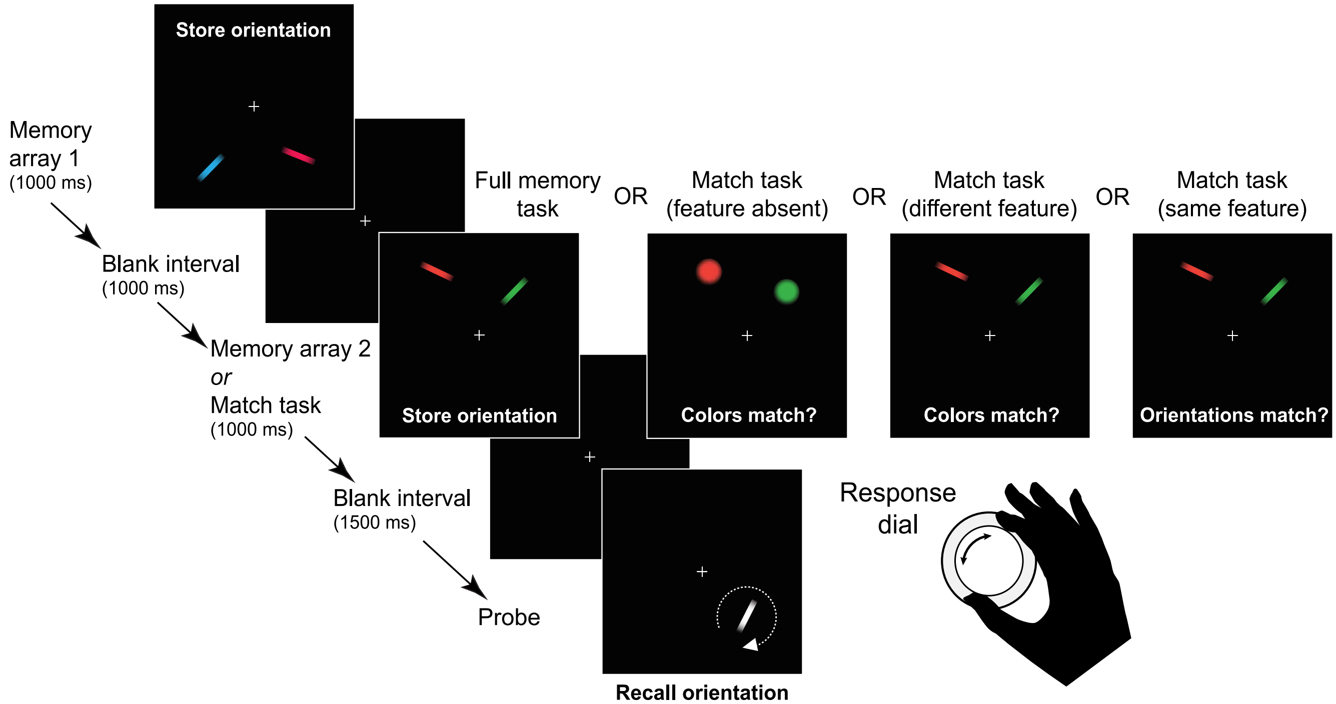

Figure 3. The recall task used in Experiment 2A. Each trial in Experiment 2A began with presentation of a memory array consisting of two colored and oriented bars. Subjects were instructed to remember the orientations of the bars. After a blank interval, a second array was presented. In the *full-memory* condition, this was a second memory array consisting of two additional oriented bars which subjects also had to store in memory. In the remaining three conditions, subjects did not have to remember the second display, but instead performed a feature-matching task based on its contents. In the *match-feature-absent condition*, subjects compared the colors of two circles and made a response if they matched. In the *match-different-feature condition*, subjects compared the colors of two oriented bars. In the *match-same-feature* condition, subjects compared the orientations of two bars. Auditory feedback was given immediately following the matching task. At the end of each trial, a probe bar was displayed, and subjects used a response dial to adjust the orientation of the probe to reproduce the remembered orientation of the item that had appeared in the same location in the preceding displays (always the first display in *match* conditions). The procedure was identical in Experiment 2B, except that the order of presentation of memory and matching arrays was reversed.

assessed the precision with which subjects could recall the orientation of bars under four different conditions, presented in separate blocks in a counterbalanced order. The procedure is illustrated in Figure 3.

In all conditions, the first display consisted of two randomly-colored and oriented bars. Subjects were instructed to memorize the orientations of the two bars. After a blank interval, a second display was presented, which differed according to experimental condition. In the *full-memory* condition, this second display was a second memory array consisting of a further two randomly-colored and oriented bars. In this condition, subjects could be asked to recall the orientation of any one of the four bars presented across the two displays, with equal probability.

In the remaining conditions, subjects were only ever required to remember the orientations of the bars in the first display. However, subjects had to attend to, and make a comparison between, specific features of objects presented in the second array. In the *match-feature-absent* condition, the second array comprised two colored circles and subjects were instructed to make a speeded button press if the circles matched in color, and to withhold a response if the colors differed. In the *match-different-feature* condition, the second array comprised two colored, oriented bars, and subjects were required to press the button only if the bars matched in color. In the *match-same-feature* condition, the second array comprised two colored, oriented bars, and subjects were required to press the button only if the orientations of the two bars matched.

The task-relevant features matched on 50% of trials and differed (by at least 45° in the space of colors/orientations) on the remaining 50%. On each trial, subjects were provided with auditory feedback immediately following the second array. If they had responded correctly, they heard a high tone, whereas if they had made the incorrect response, they heard a low tone. Following the second array, subjects were tested on their recall of the orientation of one of the bars presented in the first array.

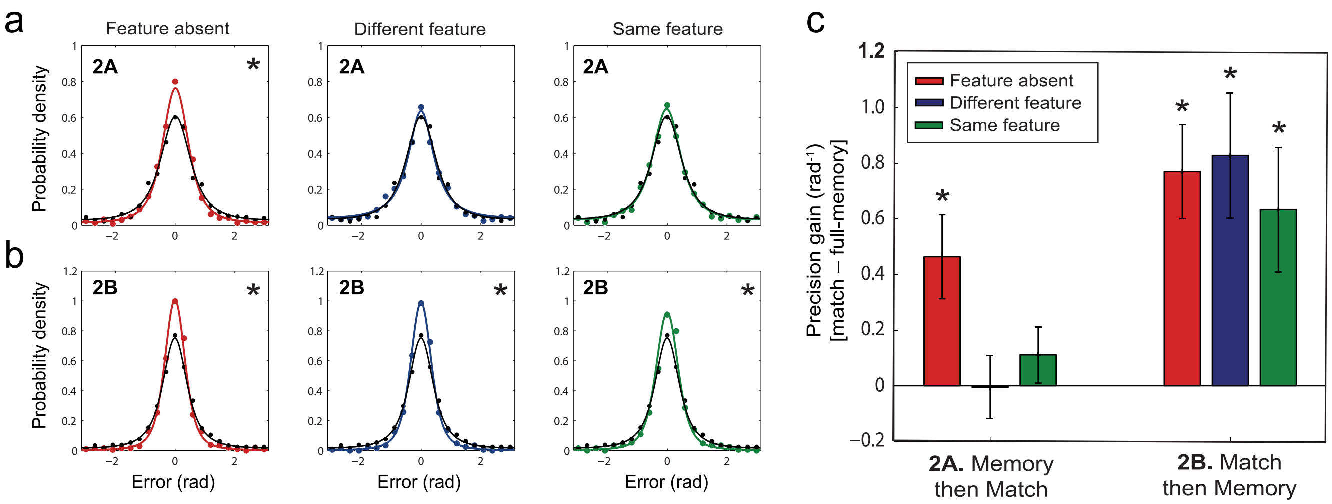

Figure 4. Error distributions and recall precision in Experiment 2. (a) Distribution of errors in orientation recall in the *match-feature-absent* (red), *match-different-feature* (blue) and *match-same-feature* (green) conditions in Experiment 2A (memory array precedes matching task). Error distributions for the corresponding trials in the *full-memory* condition are shown in black. Asterisks indicate conditions where precision was significantly enhanced relative to the *full-memory* condition ($p < 0.05$). Data points indicate mean probability over subjects. Curves correspond to subject-averaged best-fitting wrapped stable distributions (see Methods). (b) Distribution of errors in the same conditions in Experiment 2B (matching task precedes memory array). (c) Mean difference in recall precision (*precision gain*) between each of the *match* conditions and the *full-memory* condition in Experiment 2A (left) and 2B (right). Error bars indicate ±1 *SE*. Asterisks indicate significant precision gains ($p < 0.05$).

Before starting each new block of trials, subjects completed 10 practice trials to familiarize them with the displays and instructions specific to the condition. These practice trials were discarded from analysis. Subjects completed 70 trials in each of the *match* conditions, and 140 trials in the *full-memory* condition (of which 70 trials tested memory for the first array and 70 the second). Item locations were randomly selected using the same procedure as Experiment 1.

In Experiment 2B we examined the effect of reversing the order in which the memory and match arrays were presented. 10 subjects participated (six male, four female; aged 18-36 years). The procedure was identical to that in Experiment 2A except that, in the match conditions, subjects had to compare the colors or orientations of items in the first array and memorize the orientations in the second array.

Analysis

Recall precision was calculated separately for each subject and experimental condition in Experiments 2A and 2B. To assess whether subjects were successful in attending to features of an object without allocating memory resources to them, we compared precision in each of the three *match* conditions (when subjects only had to remember two orientations) to precision in the *full-memory* condition (when subjects had to remember four orientations). Specifically, for each subject, we subtracted from the precision measures obtained in each of the *match* conditions the precision calculated for items in the same display (first or second memory array) in the *full-memory* condition. This provided a measure of the *precision gain* associated with reducing the number of features to be remembered in each of the *match* conditions.

Results and discussion

All subjects were successful at performing the "matching" tasks, with mean proportion of correct responses in the range 84%–92% for all match conditions in Experiments 2A and 2B.

Figure 4a illustrates the distribution of orientation recall errors in each of the match conditions in Experiment 2A (red, blue, green), along with errors observed in the corresponding full-memory condition (black). A repeated-measures ANOVA on the precision gain compared to the full-memory condition revealed a significant overall effect of matching condition, $F(2, 22) = 7.5$, $p = 0.003$.

In the match-feature-absent condition of Experiment 2A (red), the matching task required subjects to attend to the colors of non-oriented stimuli (circles) while maintaining orientation information in memory. Recall precision was higher in this condition than in the *full-memory* condition, where twice as many orientations had to be maintained in memory, $t(11) = 3.1$, $p = 0.011$ (Figure 4c, left). This is consistent with the previously-observed relationship between increasing memory load and decreasing precision of recall.

By contrast, in the *match-different-feature* (blue) and *match-same-feature* (green) conditions, the matching tasks required subjects to attend to oriented stimuli (bars) while maintaining orientation information in memory. The two conditions differed as to whether the feature dimension that was relevant to the matching task was the same as the one held in memory (orientation) or different (color). In both cases, recall precision was indistinguishable from that observed in the *full-memory* condition, $t(11) < 1.2$, $p > 0.28$. This indicates that the requirement merely to attend to oriented objects in the matching task was responsible for an increase in the variability of orientation recall comparable to that observed in the *full-memory* task when twice the number of orientations had to be stored. This was the case even when the feature dimension of relevance to the matching task was different to the one held in memory (*match-different-feature* condition).

Attending to items in the matching array resulted in the same cost to the precision of objects already in memory as if subjects had been required to memorize those items. This result implies that the matching task required objects to be encoded to a level where they competed with other items for working memory resources. Having established that the objects were encoded, in Experiment 2B we aimed to test whether the attended objects were also involuntary *maintained* in VWM, or whether, having been encoded, they could be "dropped" from memory to make room for new information. We therefore reversed the order of the matching and memory arrays relative to Experiment 2A, such that the task requiring attention to a visual array preceded the task requiring storage in memory.

Figure 4b illustrates the distribution of orientation recall errors in each of the match conditions in Experiment 2B (red, blue, green), along with errors observed in the corresponding full-memory condition (black). Recall precision was significantly higher in all three matching conditions than in the *full-memory* condition, $t(9) > 2.8$, $p < 0.02$ (Figure 4c, right). A repeated-measures ANOVA indicated that this precision gain was equivalent across the three matching conditions ($F(2, 18) = 0.57$, $p = 0.57$). So, unlike in Experiment 2A, we found no evidence for a precision cost associated with attending to oriented visual objects immediately *prior* to presentation of a memory array. We can therefore conclude that any VWM resources allocated to the objects in the first array as a result of the matching task were successfully reallocated to the objects presented in the subsequent memory array, which were remembered with the same precision as when no competing orientation information was presented.

As in Experiment 1, we examined how well response errors could be described by two different distributions: a mixture distribution comprising uniform and Von Mises components, and a single wrapped stable distribution with variable kurtosis. The wrapped stable distribution provided a substantially better fit to the data (relative BIC, 45.4). Curves plotted in Figures 4a and b correspond to the best-fitting wrapped stable distributions for each condition.

We examined the distribution of responses relative to non-target feature values for evidence of non-target responses. Consistent with results from Experiment 1, we found a strong trend for central tendency in this distribution, although in this case it fell just short of the threshold for statistical significance (V test, $p = 0.055$). We found no significant differences in non-target error distributions between the different conditions of Experiment 2A and B (Kolmogorov-Smirnov test, $p > 0.45$), indicating that changes in the frequency of non-target responses are unlikely to have contributed substantially to the results.

General discussion

In this study we used a novel methodology to investigate the allocation of working memory resources to visual objects presented across multiple displays. Taking advantage of the established relationship between memory load and variability, we used recall precision as a probe of the contents of VWM. We found evidence for involuntary encoding of task-irrelevant features when individuals were required either to memorize a different feature of the same object (Experiment 1) or attend to a new object that was irrelevant to the memory task (Experiment 2A). Recall precision was indistinguishable in these conditions from a task in which participants were explicitly required to retain all the features in memory for a subsequent test. Importantly, when the task-irrelevant features were removed from the displays, significant gains in recall precision were observed, indicating that only in this situation was there a reduction in the number of features entering VWM.

Several previous studies have taken different approaches to investigate the encoding of task-irrelevant features, with conflicting results. Woodman and Vogel (2008) examined participants' ability to detect changes to briefly-presented, masked arrays of multi-feature items in conditions in which only color, only orientation, or a conjunction of both feature dimensions were task-relevant. Change detection performance increased with exposure duration for all conditions, but performance was superior in color than in orientation or conjunction conditions. The authors interpreted this result as indicating that consolidation of color information was slowed when orientation information also had to be stored, implying that the orientation information could be selectively excluded from encoding when it was irrelevant. However, the three conditions in this previous study differed not only in the information that needed to be remembered, but also in the comparisons participants had to make with the test array. An alternative interpretation of these results is that detection of orientation changes was simply more difficult than of color changes in this task, and that performance achieved in the conjunction condition was limited by the more difficult of the two comparisons.

In the present study, we used a reproduction rather than change detection task, so the judgment required at test was identical for all conditions. In contrast with Woodman and Vogel (2008), we observed no performance advantage in conditions where an item's color or orientation was made task-irrelevant, compared to conditions where both features were relevant. Because we used relatively long, unmasked exposures, we cannot rule out the possibility that the *rate* of encoding differed between these conditions; however, our results indicate that the outcome of processing was the same in each case, i.e., both task-relevant and task-irrelevant features were encoded into VWM. Furthermore, the use of longer exposures makes it difficult to discount this result on the grounds that selection processes had insufficient time to operate, or that encoding was still on-going at the time of array offset (Bays et al., 2009; Bays et al., 2011a).

Woodman and Vogel (2008) also examined the contralateral delay activity (CDA), a lateralized event-related potential (ERP) component associated with VWM maintenance (Van Dijk, Van der Werf, Mazaheri, Medendorp, & Jensen, 2010; Vogel, McCollough, & Machizawa, 2005; Vogel & Machizawa, 2004). Lower CDA was observed in the color-only condition than in the conjunction condition, which could be interpreted as reflecting incomplete encoding of orientation information in the condition in which it was irrelevant to the task. However, a corresponding effect was not observed in the orientation-only condition, in which color was task-irrelevant, but CDA amplitude was found not to differ from the conjunction condition. A subsequent study by Luria and Vogel (2011) found evidence that CDA amplitude was sensitive to the number of objects but not to the number of features in a memory array, which would make the CDA unsuitable as a probe of within-object feature selectivity.

Luria and Vogel argued that the relationship they observed between CDA and total number of objects supports the view that multiple features of an object are stored together in an integrated "whole object" representation. This integrated-object account is consistent with the present results; however, the interpretation of the CDA as a simple index of the contents of memory is complicated by studies showing that the CDA is also modulated by object complexity (Luria, Sessa, Gotler, Joliceur, & Dell'Acqua, 2010) and storage resolution (Machizawa, Goh, & Driver, 2012).

A recent fMRI study (Xu, 2010) found evidence for encoding of task-irrelevant object features in an analysis of BOLD signal changes recorded in two regions associated with working memory activity: the superior intraparietal sulcus (IPS) and the lateral occipital complex (LOC). The number of unique shapes and colors was varied in a task in which only memory for color was task-relevant. While the BOLD signal in superior IPS was found to be sensitive only to the number of colors, the signal in LOC depended both on the number of unique colors and on the number of unique shapes, despite the latter's irrelevance to the memory task.

The effects of increasing feature load in LOC decayed more rapidly than those in superior IPS, leading the author to propose that, although initial encoding of task-irrelevant features may be automatic, maintenance of this information is under voluntary control. This conclusion is supported by the results of our Experiment 2B, in which we reversed the order of attentional and memory tasks. When the attention-demanding matching task was presented second (Experiment 2A), the presence of competing visual features reduced precision for items already stored in memory, consistent with automatic encoding of task-irrelevant features of the matching array into VWM. However, when the matching task was presented first (Experiment 2B) no precision cost was observed for subsequently-presented material, suggesting that the information encoded from the matching task could be voluntarily forgotten, or intentionally overwritten by information from the subsequent memory array.

This observation is consistent with the established finding that informative cues presented during the maintenance of visual information ("retro-cues") can produce robust advantages for the cued item on subsequent recall (Griffin & Nobre, 2003; Kuo, Stokes, & Nobre, 2011; Makovsik & Jiang, 2007). These previous results, based on change detection tasks, suggest that voluntary control can be exerted over maintenance of items already in memory. A recent study by Pertzov, Bays, Joseph, and Husain (2012) examined the effects of probabilistic retro-cues on the precision of items in working memory. They observed that recall precision declined the longer items were maintained in memory, but that memory for an item indicated by a valid retro-cue was relatively protected from decay. This advantage came at the cost of a more rapid decline in precision for uncued items, suggesting that the retro-cue triggered a partial withdrawal of memory resources from the uncued items to support continuing maintenance of the cued item.

The present results suggest that object- or location-based attention leads to all of an object's features being encoded into VWM. However it does not necessarily follow that these features are maintained together in a single integrated object-file. Two recent studies have examined participants' ability to recall two features of one object from a multi-object array (Bays et al., 2011b; Fougnie & Alvarez, 2011). Both studies found that errors occurred with independent frequency in each feature dimension. This result presents a challenge to models in which objects are stored and fail as integrated units (Cowan, 2001; Luck & Vogel, 1997; Zhang & Luck, 2008, 2009), but it is not inconsistent with the present results. Our current findings demonstrate that participants are unable to voluntarily withhold encoding of an attended object's features into VWM, but they do not preclude the possibility of involuntary failure or variability in the encoding or maintenance process for each feature, resulting in independent errors on subsequent recall.

The distribution and source of errors in recall tasks has become a focus of considerable research and debate in recent years (Anderson et al., 2011; Bays et al., 2009; Bays et al., 2011a; Bays & Husain, 2008; Bays et al., 2011b; Fougnie et al., 2010; Fougnie & Alvarez, 2011; Gorgoraptis et al., 2011; Lara & Wallis, 2012; Murray, Nobre, Astle, & Stokes, 2012; Van den Berg et al., 2012; Wilken & Ma, 2004; c & Luck, 2008, 2009). While it was not the main goal of the present study to discriminate between competing models of error distribution, we would briefly note two observations. First, consistent with previous studies (Bays et al., 2009; Bays et al., 2011b), we observed a small but significant contribution to performance of errors centered on the feature values of non-target (unprobed) items in the memory array. These errors occur when a participant mistakes which of the items in memory was at the probed location, because of either variability in memory for location or errors in maintaining the binding information that links features together. Importantly, the frequency of these errors did not vary between the different conditions of our experiments, so they are unlikely to have contributed to the main effects observed in the study.

Second, while the distribution of errors on the reproduction task deviated from Von Mises (the circular equivalent of the Gaussian distribution), it was better described by a single continuous distribution with heavier tails than the Von Mises, than by a mixture of two distributions, one Von Mises and one uniform. This result is consistent with models that describe VWM as a continuously-divisible resource (Bays & Husain, 2008; Gorgoraptis et al., 2011; Huang, 2010; Lara & Wallis, 2012; Van den Berg et al., 2012), rather than a quantized commodity (Anderson et al., 2011; Zhang & Luck, 2008). Previous results have shown that VWM resources are not always distributed equally between stimuli, and if the task requires it, can be flexibly allocated to store certain objects with higher resolution than others (Bays et al., 2011a; Bays & Husain, 2008; Lara & Wallis, 2012). Variability in the allocation of resources to objects provides one plausible explanation for the shape of the error distribution observed here (Van den Berg et al., 2012).

Importantly for the present study, our methodology does not depend on the shape of the error distribution, but only on the relatively uncontroversial observation that recall precision declines monotonically with increasing memory load, allowing us to infer decreases in memory load from precision gains. The results support a relatively inflexible view of memory encoding, in which task-irrelevant features of memorized or attended objects automatically access VWM, depleting the resources available for other objects. However, we have found evidence to suggest that this involuntary encoding is compensated by voluntary control over maintenance, allowing unwanted information to be "dropped" from memory to release resources for new objects.

Keywords: working memory, resources, attention, feature selectivity

Acknowledgment

This research was supported by the Wellcome Trust.

Commercial relationships: none
Corresponding author: Paul M Bays.
Email: p.bays@ucl.ac.uk.
Address: Sobell Department of Motor Neuroscience and Movement Disorders, UCL Institute of Neurology, London, UK.

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.

Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in Visual Working Memory Reaches a Stable Plateau When Individual Item Limits Are Exceeded. *Journal of Neuroscience, 31,* 1128–1138.

Arthur, P. (2008). The wrapped stable family of distributions as a flexible model for circular data.

Computational Statistics & Data Analysis, 52, 1516–1523.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10): 7, 1–11, http://www.journalofvision.org/content/9/10/7, doi:10.1167/9.10.7. [PubMed] [Article].

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011a). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision, 11*(10): 6, 1–15, http://www.journalofvision.org/content/10/11/6, doi:10.1167/11.10.6. [PubMed] [Article].

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321,* 851–854.

Bays, P. M., Wu, E. Y., & Husain, M. (2011b). Storage and binding of object features in visual working memory. *Neuropsychologia, 49,* 1622–1631.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Science, 24,* 87–114.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501–517.

Elmore, L. C., Ma, W. J., Magnotti, J. F., Leising, K. J., Passaro, A. D., Katz, J. S., et al. (2011). Visual short-term memory compared in rhesus monkeys and humans. *Current Biology, 21,* 975–979.

Fisher, N. I. (1995). *Statistical analysis of circular data.* Cambridge, UK: Cambridge University Press.

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision, 11*(12):3, 1–12, http://www.journalofvision.org/content/11/12/3, doi:10.1167/11.12.3. [PubMed] [Article].

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision, 10*(12):27, 1–11, http://www.journalofvision.org/content/10/12/27, doi:10.1167/10.12.27. [PubMed] [Article].

Gorgoraptis, N., Catalao, R. F., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *The Journal of Neuroscience, 31,* 8502.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience, 15,* 1176–1194.

Huang, L. (2010). Visual working memory is better characterized as a distributed resource rather than discrete slots. *Journal of Vision, 10*(14):8, 1–8, http://www.journalofvision.org/content/10/14/8, doi:10.1167/10.14.8. [PubMed] [Article].

Kuo, B.-C., Stokes, M. G., & Nobre, A. C. (2011). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience, 24,* 51–60.

Lakha, L., & Wright, M. J. (2004). Capacity limitations of visual memory in two-interval comparison of Gabor arrays. *Vision research, 44,* 1707–1716.

Lara, A. H., & Wallis, J. D. (2012). Capacity and precision in an animal model of visual short-term memory. *Journal of Vision, 12*(3):13, 1–12, http://www.journalofvision.org/content/12/3/13, doi:10.1167/12.3.13. [PubMed] [Article].

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

Luria, R., Sessa, P., Gotler, A., Joliceur, P., & Dell'Acqua, R. (2010). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience, 22,* 496–512.

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49,* 1632–1639.

Machizawa, M. G., Goh, C. C. W., & Driver, J. (2012). Human visual short-term memory precision can be varied at will when the number of retained items is low. *Psychological Science, 23,* 554–559.

Makovsik, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin and Review, 14,* 1072–1078.

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology, 14,* 744–751.

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences, 29,* 317–322.

Murray, A. M., Nobre, A. C., Astle, D. E., & Stokes, M. G. (2012). Lacking control over the trade-off between quality and quantity in visual short-term memory. *PLoS ONE, 7,* e41223.

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401,* 584–587.

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

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 332–350.

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

Pertzov, Y., Bays, P. M., Joseph, S., & Husain, M. (2012). Rapid forgetting prevented by retrospective attention cues. *Journal of Experimental Psychology: Human Perception and Performance*. Online First Publication, December 17, 2012, doi:10.1037/a00309.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision Research, 43,* 629–637.

Sperling, G. (1960). The information available in brief presentations. *Psychological Monographs, 74*(11), 1–29.

Van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, USA, 109,* 8780–8785.

Van Dijk, H., Van der Werf, J., Mazaheri, A., Medendorp, W. P., & Jensen, O. (2010). Modulations in oscillatory activity with amplitude asymmetry can produce cognitively relevant event-related responses. *Proceedings of the National Academy of Sciences, USA, 107,* 900–905.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123,* 146.

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.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*(12):11, 1120–1135, http://www.journalofvision.org/content/4/12/11, doi:10.1167/4.12.11. [PubMed] [Article].

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. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 458–468.

Xu, Y. (2010). The neural fate of task-irrelevant features in object-based processing. *The Journal of Neuroscience, 30*, 14020–14028.

Zar, J. H. (2010). *Biostatistical analysis*. Upper Saddle River, NJ: Prentice Hall.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453,* 233–235.

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20,* 423–428.