

Memory for a single object has differently variable precisions for relevant and irrelevant features

Garrett Swan

Department of Psychology, Pennsylvania State University,
University Park, PA, USA



John Collins

Department of Physics, Pennsylvania State University,
University Park, PA, USA



Brad Wyble

Department of Psychology, Pennsylvania State University,
University Park, PA, USA



Working memory is a limited resource. To further characterize its limitations, it is vital to understand exactly what is encoded about a visual object beyond the “relevant” features probed in a particular task. We measured the memory quality of a task-irrelevant feature of an attended object by coupling a delayed estimation task with a surprise test. Participants were presented with a single colored arrow and were asked to retrieve just its color for the first half of the experiment before unexpectedly being asked to report its direction. Mixture modeling of the data revealed that participants had highly variable precision on the surprise test, indicating a coarse-grained memory for the irrelevant feature. Following the surprise test, all participants could precisely recall the arrow’s direction; however, this improvement in direction memory came at a cost in precision for color memory even though only a single object was being remembered. We attribute these findings to varying levels of attention to different features during memory encoding.

Introduction

Human working memory is resource limited, and understanding the nature of these limits is a major goal of working memory research (Baddeley, 2003; Brady, Konkle, & Alvarez, 2011; Luck & Vogel, 2013; Ma, Husain, & Bays, 2014). Quantitatively exploring human memory performance by measuring the limits on capacity and precision provides important clues about the underlying neural mechanisms of memory storage. One measured limit has been characterized as the number of simple objects (about three to four; Cowan, 2001) that can be stored in visual working

memory.¹ Furthermore, when a task involves attending to particular objects in a scene, it can happen that task-irrelevant objects are not remembered at all, even when they are salient (Rock, Linnett, Grant, & Mack, 1992; Simons & Chabris, 1999). When task-irrelevant objects do attract attention, this often comes at the expense of maintaining task-relevant information (Asplund, Todd, Snyder, Gilbert, & Marois, 2010; Horstmann, 2005, 2006; Yin et al., 2012). Additionally, information that was once task relevant does not necessarily remain in memory (Chen & Wyble, 2015; Triesch, Ballard, Hayhoe, & Sullivan, 2003). These findings demonstrate natural strategies for efficiently using limited memory resources.

We often perceive the world without an explicit expectation of what information we will need to remember for later use, and very little is known about how task-irrelevant features of attended objects are represented in memory. Even the simplest objects contain many features such as size, luminance, perceived depth, hue, texture, and shape. Therefore, given the limited capacity of memory, it is important to determine how these various features of an attended object are stored in memory, how precisely they are stored, and how their storage depends on task relevance. This understanding ultimately is critical for building better models of working memory.

Multiple hypotheses can be made about the representation of task-irrelevant features. One straightforward possibility is that objects are encoded holistically, such that if an object is encoded, then all of its features are fully represented (Luck & Vogel, 1997; Zhang & Luck, 2008; see A in Figure 1). Another straightforward possibility is that only task-relevant information is represented, leading to a highly efficient representa-

Citation: Swan, G., Collins, J., & Wyble, B. (2016). Memory for a single object has differently variable precisions for relevant and irrelevant features. *Journal of Vision*, 16(3):32, 1–12, doi:10.1167/16.3.32.

doi: 10.1167/16.3.32

Received July 23, 2015; published February 24, 2016

ISSN 1534-7362



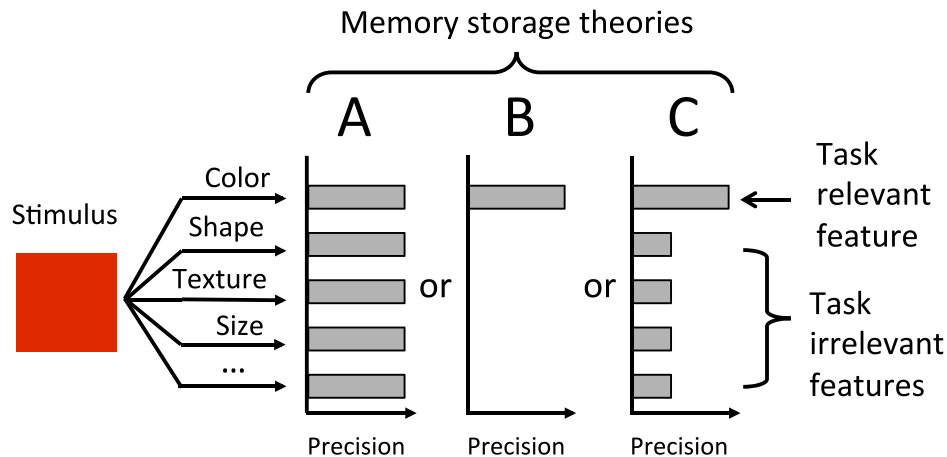


Figure 1. Illustration of how well relevant (i.e., color) and irrelevant features could be encoded into memory according to three theories. (A) refers to memory storage in which all features are fully represented. (B) refers to the memory storage of only the task-relevant feature. (C) refers to the full representation of relevant features, with coarse coding of task-irrelevant features.

tion for that feature but complete amnesia for task-irrelevant information (Awh, Vogel, & Oh, 2006; Chun, 2011; Serences, Ester, Vogel, & Awh, 2009; Woodman & Vogel, 2008; see B in Figure 1).

If all features were fully represented, then performance on tasks with a variable number of features per object would depend only on the number of objects, as found by Luck and Vogel (1997). However, several findings have challenged this result (e.g., Chen & Wyble, 2015; Fournie, Asplund, & Marois, 2010; Hardman & Cowan, 2015; Keshvari, van den Berg, & Ma, 2013; Oberauer & Eichenberger, 2013; Wheeler & Treisman, 2002). Furthermore, it has been demonstrated using delayed estimation that for a given stimulus with multiple relevant features, one feature might be well remembered while another is poorly remembered (Bays, Wu, & Husain, 2011; Fournie & Alvarez, 2011). Thus, the strong version of the holistic encoding theory seems unlikely.

On the other hand, the hypothesis that memory represents only task-relevant information is also unlikely because task-irrelevant features can influence reaction time (T. Gao, Gao, Li, Sun, & Shen, 2011; Kahneman, Treisman, & Gibbs, 1992; Stroop, 1935) and memory precision (Marshall & Bays, 2013). Furthermore, task-irrelevant features of objects have been shown to elicit neurophysiological responses (Z. Gao et al., 2009; Xu, 2010).

With both straightforward hypotheses unlikely, we propose that task-irrelevant features may sometimes be encoded coarsely—that is, at lower levels of precision than relevant features (see C in Figure 1). Coarse coding would enable memory to focus its limited resources on the task-relevant features of an object but also have sufficient memory for task-irrelevant features to support unanticipated task requirements.

In this article, we present direct measurements of the quality of a task-irrelevant feature's representation and of the cost of adding an extra task-relevant feature on memory precision of another feature. We used a display containing only a single object to ensure that memory load was below the typically measured memory load for objects. To test memory for task-irrelevant features, we used the surprise test methodology of Rock et al. (1992), which unexpectedly asks the participant to recall a piece of information that was encountered while that information was irrelevant to the task. Recently, Eitam, Yeshurun, and Hassan (2013) used this technique and found that participants produced more errors when asked to recognize a task-irrelevant color of a two-colored stimulus. Eitam, Shoval, and Yeshurun (2015) later found that participants could recognize the irrelevant color of a single-colored stimulus, which supports the holistic encoding hypothesis. However, good performance in such forced-choice tasks does not necessarily mean that the memory is precise. It could be that the memory for an irrelevant color is coarsely coded but still sufficiently precise to select the correct color from highly dissimilar colors in a forced-choice response.

To measure the precision of memory for task-irrelevant features, we paired the surprise test methodology with delayed estimation, which requires participants to reconstruct a specific feature value of a presented stimulus by selecting that feature along a continuous scale (Prinzmetal, Amiri, Allen, & Edwards, 1998; Wilken & Ma, 2004). This technique provides within-trial measurements of memory error relative to the true value of the stimulus. Furthermore, to obtain the cleanest possible measure of the effect of relevance, we used memory for a single object with clearly dissociable features, and we trained participants on

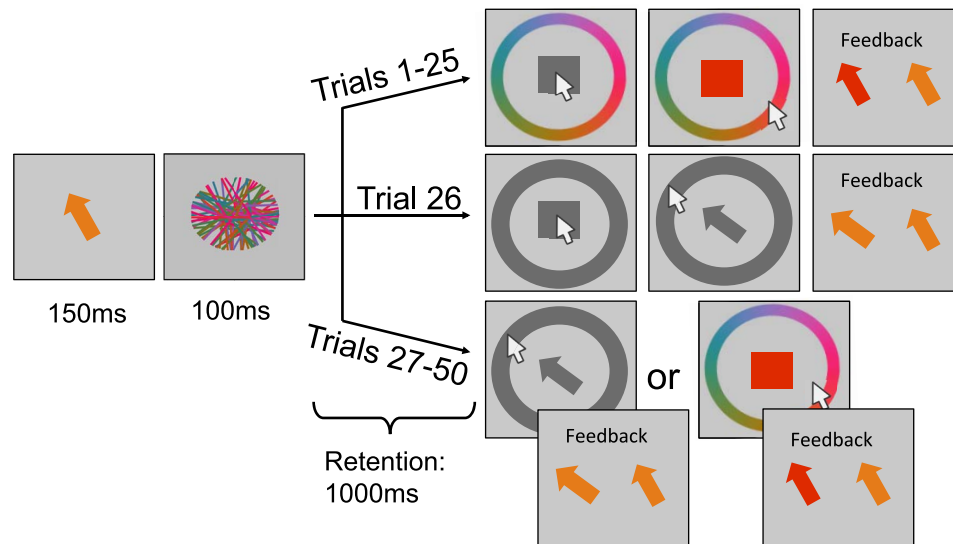


Figure 2. Illustration of the task. Participants were presented with a colored arrow for 150 ms. After the offset of the arrow, a mask comprising random color and oriented lines was presented for 100 ms. After a retention interval of 1000 ms in which the screen was empty, participants responded either by selecting a hue from a color wheel or by selecting a direction using a gray wheel. After providing responses, participants were given feedback about their response. The color wheel and task were generated using modifications of MemToolbox (Suchow et al., 2013).

reporting the task-irrelevant feature prior to the start of the experiment.

Our study used displays with a colored arrow in which we manipulated color and direction (Figure 2). One of these features (color) remained task relevant throughout the experiment, while the other feature (direction) unexpectedly switched from irrelevant to relevant halfway through the experiment. This provided an explicit measurement of the quality of the memory for the initially task-irrelevant feature the first time it was probed, which we termed the *surprise trial*. The postsurprise trials then measured the cost on the always-relevant feature when the other feature became relevant halfway through the experiment.

Experiment design and analysis

Participants

To obtain a sample of 150 participants after replacing excluded participants, a total of 175 participants were recruited from the participant pool at the Pennsylvania State University and received course credit for their participation. All participants had normal or corrected-to-normal vision and could read American English. The experiments conformed to the Declaration of Helsinki and were approved by the Pennsylvania State University Institutional Review Board ethics committee. A total of 22 participants were

excluded for having poor perceptual matching accuracy (see below), and three participants were excluded for having poor accuracy in the second block of trials (one with postsurprise color $\sigma = 63$, one with more than 40% guess rate postsurprise direction, and one with postsurprise direction $\sigma = 83$). Note that the inclusion of the 25 excluded participants does not affect the significance of the cost between pre- and postsurprise color precision and the results of the model fitting to the surprise trial data.

Apparatus

The experiment was run using Matlab 7.9.0 (build R2009b) with Psychtoolbox (Brainard, 1997; Pelli, 1997) on Windows XP (Microsoft, Redmond, WA). The screen resolution was set to 1024×768 at a 75-Hz refresh rate on cathode ray tube monitors with a diagonal screen size of 40 cm. Participants were situated in chin rests located 50 cm from the monitor.

Stimuli

The stimulus consisted of an arrow with a direction and color feature. The arrow was constructed using a rectangle with the dimensions 5.4° and 1.2° of visual angle and an isosceles triangle with a 1.7° base and 1.2° height, which combined to form an arrow with a height of 6.2° . Direction was chosen from 360° as indicated by the arrowhead. The color of the arrow was drawn from

a series of 360 colors in a one-dimensional selection from CIE $L^*a^*b^*$ color space provided by Memtool-Box (Suchow, Brady, Fougne, & Alvarez, 2013). The arrow was centered at fixation during the experimental trials.

Perceptual matching

To ensure that participants knew how to produce estimates of color and direction, a perceptual matching task preceded the experimental trials. This block was labeled *Experiment 1*. In this task, two stimuli appeared along the horizontal meridian separated by 10.8° of visual angle. Participants were instructed to match a single feature of the left stimulus (i.e., patch for color matching and arrow for direction matching) to the right stimulus by selecting a position on the surrounding circle with a mouse click. The colors and directions of the right stimulus were discretely sampled (i.e., 36° difference) to cover the extent of the wheel and were presented in a random order that was fixed across participants. The appearance of the left circle changed depending on which feature was being matched. The wheel was colored for color matching and solid gray for direction matching. This difference in color wheel presentation was also used during the delayed estimation portion of the experiment. When participants moved their mouse to the surrounding wheel, the feature of the stimulus changed to represent the appropriate feature (e.g., the patch became the color associated with the location of the mouse for color retrieval, or the arrow pointed at the location of the mouse for direction retrieval). Participants were free to move the mouse until satisfied with their answer, and the left stimulus changed continuously to represent the direction or color indicated by the current mouse position. The orientation of the color wheel was unchanged across trials throughout the experiment for both the perceptual matching and the delayed estimation blocks.

When participants were satisfied that the stimulus on the left matched the stimulus on the right, they would click on the wheel using the left mouse button to lock in their choice. After a participant made a response, the next trial began in 500 ms. Participants completed two blocks of 10 trials. The only difference between blocks was whether participants were matching the color or direction of the stimulus.

Error was calculated by taking the difference between the reported angle and the presented angle. Participants with errors above 10° on any of the last three perceptual matching trials for reporting direction (i.e., 22 participants) were excluded to ensure that participants could understand and perform the task

requirements on the surprise trial. After exclusion, we found that the average of the absolute value of the error in matching direction over the last three trials was 2.1° . Note that participants were much more precise on this perceptual matching task than on the delayed estimation task (e.g., compare with Figure 3), which formed our main experiment; this was expected from previous work (Brady, Konkle, Gill, Oliva, & Alvarez, 2013).

Delayed estimation

After completing the perceptual matching task, participants began the experimental trials that constituted the delayed estimation task, which was labeled *Experiment 2*. Participants were instructed at the beginning of this portion of the experiment that their task was to remember the color of the arrow (see the Appendix for exact instructions presented to participants). For each trial, following a blank presentation for 300 ms, an arrow appeared for 150 ms. Immediately following the stimulus duration, a mask of 50 randomly oriented and colored lines was presented with a radius of 5.4° of visual angle for 100 ms. Next, there was a retention interval of 1000 ms, followed by the response screen. The response screen was similar to the perceptual matching task except that there was only a single stimulus in the center of the screen. After a response was made, a white line appeared at the location of the mouse click, extending outward from the wheel; the mouse pointer relocated to the center of this line. Participants were asked to report their confidence by clicking on the white line, with locations at the farthest extent of the white line indicating maximum confidence. Confidence scores were collected, although those data are not discussed here. After selecting a confidence value, feedback was provided as to the original stimulus. For the feedback screen two arrows were presented, with the arrow on the left mirroring the response of the participant and the arrow on the right being the presented arrow. Text indicated which was which. The next trial began 500 ms after a mouse click.

Participants completed six practice trials and then 50 experimental trials. The practice trials and experimental trials 1 through 25 asked participants to report only color. Trial 26 asked participants to report direction, with the following message appearing between the surrounding wheel and the central stimulus: “This is a surprise test. What was the direction of the arrow?” For the remaining trials (trials 27–50), participants were asked at random to report either color or direction.

Results

Both features were recalled by making a selection from a circle, and memory error was computed as the angular difference between the correct value (color or direction) and the reported value. A simple measure of the participants' performance is the mean of the absolute value of the errors, as shown trial by trial in Figure 3. On the surprise trial, where participants were unexpectedly asked to report the direction, their average error was quite poor. Their precision in reporting direction improved on the first postsurprise trial and then remained stable. However, this increased precision in reporting direction came at the expense of worse precision in reporting color, as is visible in Figure 3 and is quantified below and in the Appendix.

A fuller picture of the changes in recall precision is shown in Figure 4, which shows distributions of errors for color and for direction in three portions of trials: presurprise, surprise, and postsurprise. First, we observe that the postsurprise color recall and the postsurprise direction recall are peaked near zero with only a small tail. This indicates that the memory system is well within its capacity limits regarding the number of objects and features even when participants are reporting both color and direction (Bays, Catalao, & Husain, 2009; Zhang & Luck, 2008).

However, the surprise trial direction responses, for which there is one trial per participant, are qualitatively different from the other distributions. Most notably, the surprise trial data have a long tail out to the largest error values. The narrow peak in the middle suggests that at least some participants stored the task-irrelevant feature precisely. The complete data set from this experiment is analyzed for the task-

relevant feature first, followed by the task-irrelevant feature.

Memory quality for the always-relevant feature

After the surprise trial, the average error of participants' reports of direction improved dramatically. This suggests that a shift in expectation about stimulus reporting changed the participants' attentional set such that they encoded a more precise representation of the initially task-irrelevant feature. To determine whether this change in the attentional set produced a cost for memory of the arrow's color, we compared the memory quality for color between the pre- and postsurprise trial data as follows.

To allow for interparticipant variability, for each participant we fit separately the pre- and postsurprise trial distribution of errors using the two-component mixture model of Zhang and Luck (2008; henceforth referred to as the *ZL model*):

$$p(x|\kappa, P_u) = (1 - P_u)M(x|\kappa) + P_u \frac{1}{2\pi} \quad (1)$$

where the mixture model is a combination of a von Mises distribution, $M(x|\kappa)$, with a freely varying concentration parameter (κ), and a uniform distribution, which is assumed to result from guessing. The concentration parameter was converted from κ to circular standard deviation (σ) in degrees (Fisher, 1995):

$$\sigma(\kappa) = \frac{180}{\pi} \sqrt{-2 \ln \frac{I_1(\kappa)}{I_0(\kappa)}} \quad (2)$$

where $I_0(\kappa)$ and $I_1(\kappa)$ are modified Bessel functions. The width of the von Mises distribution (σ) and the

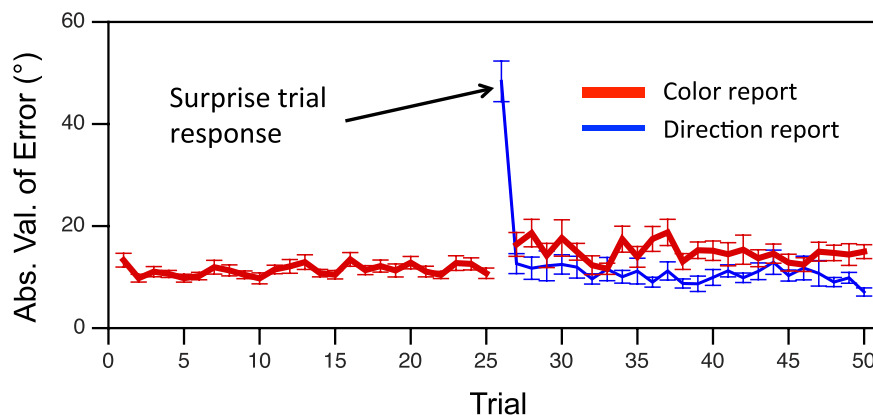


Figure 3. Each data point represents the mean absolute value of the error for all participants on a given trial. For the first 25 trials, participants were consistently asked the color of the presented color arrow. On trial 26, participants were given a surprise test, which probed their memory for the direction of the arrow. For the remaining trials, participants could be asked about either the color or the direction of the arrow. Error bars denote standard errors.

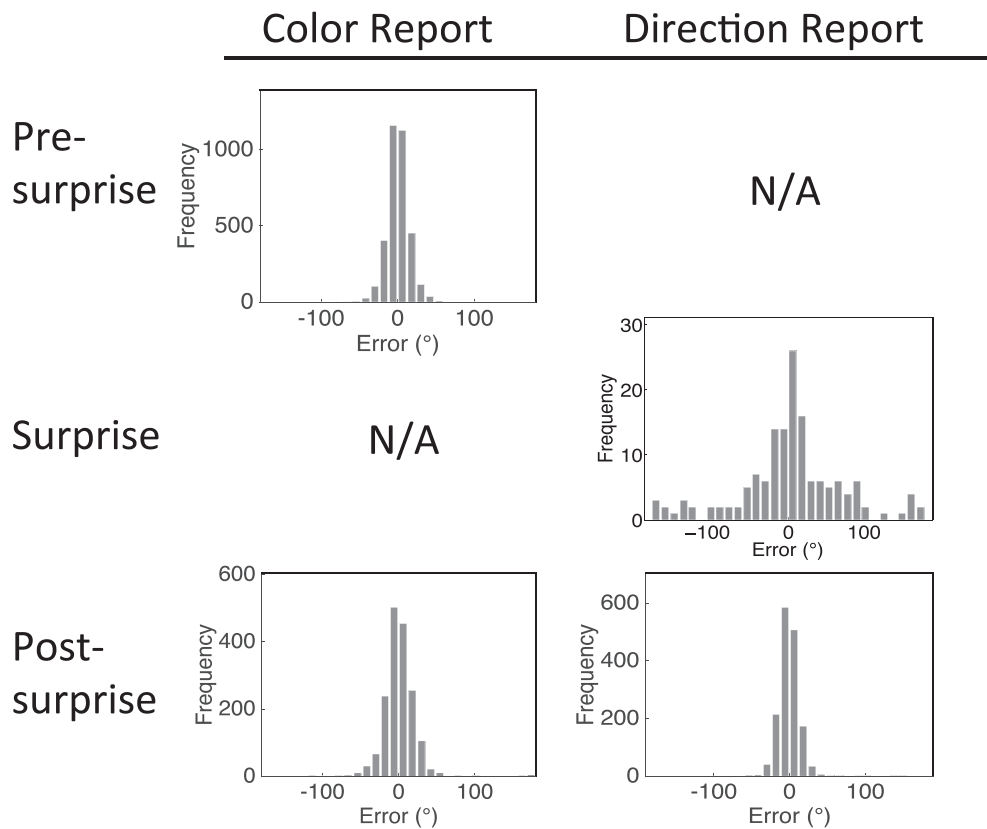


Figure 4. Histograms of error distributions across participants for both color and direction recall in presurprise, surprise, and postsurprise trials. Note that the surprise trial was always direction recall and that each count represents a single data point from a participant. Also note that it is difficult to see the small percentage of data points in the tails of the presurprise color and postsurprise color and direction distributions.

proportion of guesses (P_u) were fit to the data using maximization of the likelihood:

$$LL(\kappa, P_u) = \sum_{i=1}^M \ln(p(x_i|\kappa, P_u)) \quad (3)$$

where M refers to the total number of trials for a given participant and condition.

The proportion of guesses was quite small for presurprise ($M = 0.015$, $SEM = 0.003$) and postsurprise ($M = 0.032$, $SEM = 0.006$) color responses. This was expected given the low-load nature of the task. Similarly, participants were precise in reporting the color in both presurprise ($M = 12.9^\circ$, $SEM = 0.26$) and postsurprise ($M = 15.3^\circ$, $SEM = 0.45$; Figure 5) trials but had worse precision postsurprise.

To determine the significance of the cost in precision for color when direction became relevant, we applied a paired t test for the difference between the pre- and postsurprise fitted widths for color responses per participant: $t(149) = 5.2$, $p < 0.001$, 95% confidence interval [1.5, 3.4]. This highly significant effect indicates that color report became less precise when participants adopted the requirement to report direction as well. We

considered and eliminated the possibility that this decline in color precision was the result of a gradual decline in color precision across the entire set of 50 trials (see the Appendix). Similarly, this difference in

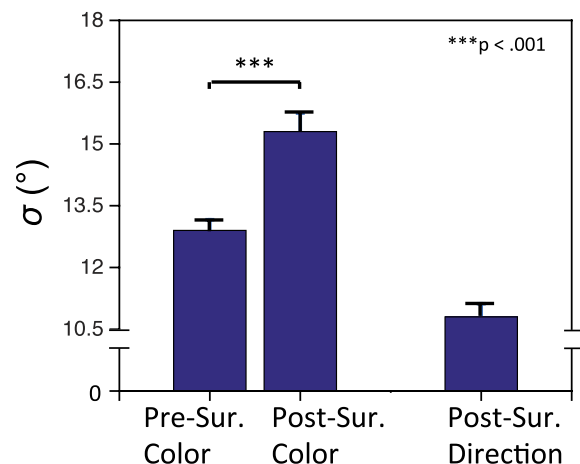


Figure 5. Average fitted standard deviation parameters across participants using the ZL model. The difference between the pre- and postsurprise color conditions represents the cost of also having to report direction in the postsurprise trials.

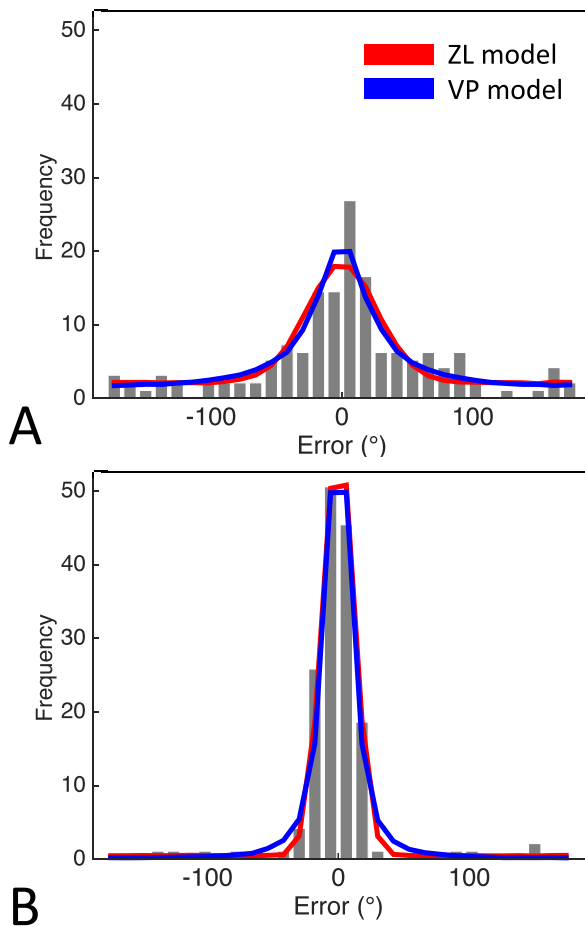


Figure 6. Histogram of responses with model fits placed onto the data. (A) ZL_s model and VP model fit to the surprise trial data across participants. (B) Fits of the ZL_s model (red) and the VP model (blue) to the first postsurprise direction trial data across participants.

precision was found in a fixed effects model comparison analysis (see the Appendix). Furthermore, there is a significant difference between the pre- and postsurprise proportion of guesses for color responses per participant, $t(149) = 2.6$, $p < 0.01$, 95% confidence interval [0.004, 0.03], although the number of guesses is small in both cases.

In an additional control experiment, the feature relevance conditions were reversed such that the color of the arrow was the initially task-irrelevant feature during the first 25 trials. When fitting the ZL model to the presurprise and postsurprise direction data, participants ($N = 30$) were less precise postsurprise ($M = 10.1^\circ$, $SEM = 0.73$) than they were presurprise ($M = 8.2$, $SEM = 0.44$). This difference was significant in a paired t test, $t(29) = 3.2$, $p < 0.005$, 95% confidence interval [0.63, 3.1], which further corroborates our conclusion that there is a cost to encoding an additional feature in a single, simple object.

Memory quality for the initially task-irrelevant feature

For the initially task-irrelevant feature (direction), we first fitted the postsurprise direction reports for each participant, again with the ZL model (Zhang & Luck, 2008), obtaining a precision averaged over participants of 10.8° ($SEM = 0.32$) and a small uniform component ($M = 0.025$, $SEM = 0.005$). These properties quantify the precision of direction memory when both direction and color were relevant.

Then we examined participants' responses for the direction of the arrow in the surprise trial, when direction was considered task irrelevant from the perspective of the participant at the time it was observed. Because there was one surprise trial per participant, we fitted the two-component ZL model to the data pooled over participants using maximum likelihood estimation, with

$$LL(\kappa, P_u) = \sum_{i=1}^N \ln(p(x_i|\kappa, P_u)) \quad (4)$$

where N refers to the total number of participants and i labels participants. We refer to this version of the ZL model that is across participants as the ZL_s model. The precision and guess rate for the responses on the surprise trial were found to be 26.1° and 0.42, respectively (Figure 6a).

To contrast the precision of direction on the surprise trial to that on the postsurprise trials, we compared the surprise trial response to the response in the first postsurprise trial in which participants reported the direction of the arrow. This is the most conservative comparison with the surprise trial and highlights how quickly direction shifted from task irrelevant to relevant. The fit for the first postsurprise direction trial using the ZL_s model produced a precision and guess rate of 12.1° and 0.07, respectively (Figure 6b).

To determine whether there is a reliable difference in the fit likelihoods for the two models (i.e., surprise and the first postsurprise ZL_s fits), a permutation over differences in log likelihood (LL) was performed. First, the LL was computed for the best fit for the surprise data to the ZL_s model. Then, the LL was computed for the same data but with the parameters obtained for the best fit to the postsurprise data (i.e., a model fit to a different data set). The difference (70.5) between these LL values is good evidence that the surprise and postsurprise data are drawn from different distributions. To produce a null distribution, the data points between these two trial types were randomly permuted 10,000 times. For each permutation, new ZL_s models were fit to each of the two shuffled trial sets and the same LL difference was computed. A comparison of the observed difference between the empirically derived

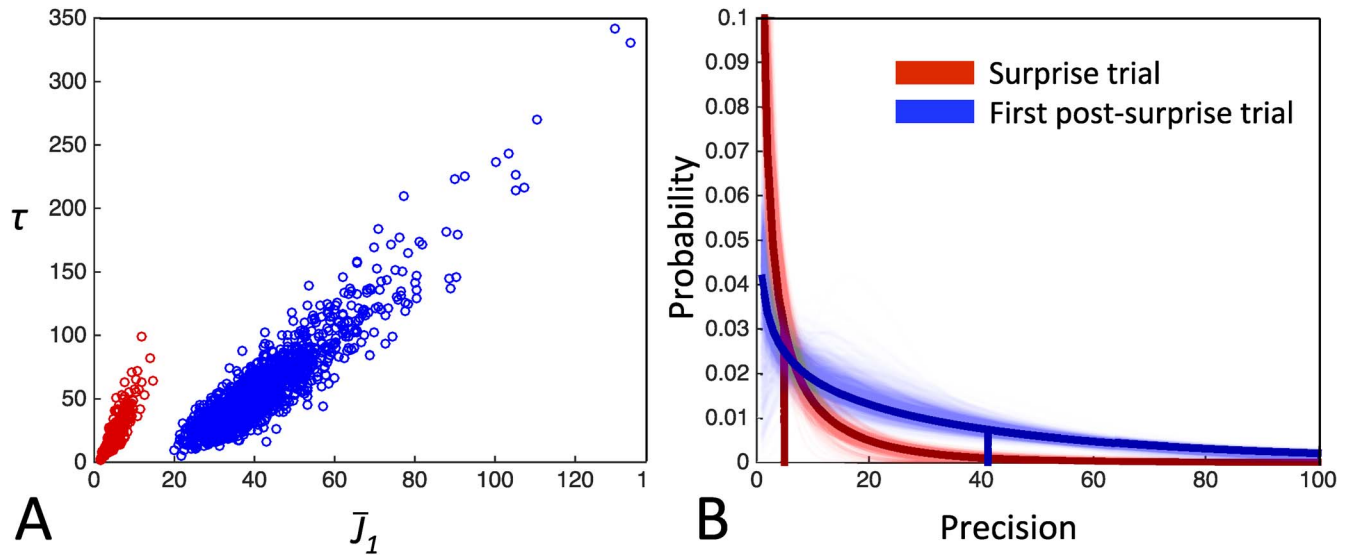


Figure 7. Results of the bootstrapped resampling 1,500 times. (A) Scatter plot of the parameters of the VP model fit. (B) Distribution of precision in the VP model. The dark lines correspond to the observed precision of the actual data, with the vertical line corresponding to the mean precision (i.e., \bar{J}).

difference and the null distribution revealed that zero of the permuted differences were larger than the observed difference for both the surprise and first postsurprise direction data, which indicates at least $p < 0.0001$.

This procedure was then repeated for the converse case (i.e., comparing the difference in LLs for when the first postsurprise trial data were fit to the first postsurprise trial data and the surprise trial data) and revealed that none of the 10,000 null distribution values were as extreme as the observed difference. These two tests demonstrate that the surprise and postsurprise distributions are substantially and significantly different, which indicates that participants are less precise on the surprise trial relative to the first postsurprise trial that probes direction.

The ZL mixture model allows for two levels of precision: high and zero (i.e., guessing). An alternative approach is to allow precision to vary continuously as in the variable precision (VP) model of van den Berg et al. (2012). In this model, the precision of a response on each trial is modeled by drawing a value from a gamma distribution with a mean precision \bar{J} and scale parameter τ . Furthermore, it is assumed that responses are affected by sensorimotor noise, which is modeled by convolution with a von Mises distribution with concentration parameter κ_r (see the Appendix for a full specification of the model).

In van den Berg et al. (2012), \bar{J} is dependent on a set size parameter, which we fixed at 1 given that only a single stimulus was presented. Furthermore, we fixed the width of the sensorimotor noise distribution κ_r at the empirically measured value determined by the perceptual matching task for the direction feature. The

value was obtained by a maximum likelihood function fit to the data on the last three trials pooled over all 175 participants using the ZL mixture model.² The value of κ_r corresponds to 2.9° . Fixing these parameters reduced the number of free parameters to two, as in the two-component ZL mixture model.

The fit of the VP model to the surprise trial data is displayed in Figure 6. In Figure 7A and B, parameters were bootstrapped by resampling 1,500 times the surprise trial data and the first postsurprise trial in which direction was reported. The difference in the parameter estimates for the surprise and first postsurprise trials is visually obvious, but to determine significance a permutation test on the mean precision parameter J was computed. In 10,000 permutations of the data, no permutations exceeded the difference of the actual data. Thus, like the ZL_s fit, the VP model shows that participants were less precise for direction on the surprise trial than on the postsurprise trials. Furthermore, the fits to the VP model for the surprise data reveal a substantial proportion of very low precision responses; this is analogous to the high guessing rate in the fit by the ZL_s model to the same data. (See the Appendix for tables of parameter values for both ZL_s and VP models.)

To see whether the VP or the ZL_s model best characterizes the data, the LL can be compared between the models because they have the same number of free parameters, having fixed the set size and sensorimotor noise parameters of the VP model (see above). The VP model produced a higher LL (LL = -215.9) than the ZL_s model (LL = -217) when fit to the surprise trial data, suggesting that the VP model provides a better fit, though the difference of 1.1 LL

units is not strong support. However, the conclusion of these analyses is similar regardless of using the VP or ZL models because both indicate a more coarsely grained representation of direction in the surprise trials.

Our conclusion, therefore, is that variability of participants' precision for the task-irrelevant feature is greatly increased on the surprise trial. Furthermore, the training procedure used at the beginning of the experiment provided reassurance that participants understood how to report direction very precisely using the mouse. Thus, the inaccuracy is not due to uncertainty about how to perform the task.

Discussion

In the field of working memory research, almost everything that is known was determined by measuring our ability to store and retrieve task-relevant information about stimuli. However, stimuli always contain additional features that are irrelevant to the task, and our understanding of how these irrelevant features are stored has been largely unexplored. The goals of our study were (a) to measure the precision of memory for an irrelevant feature and (b) to determine the costs in precision of memory for a relevant feature when a previously irrelevant feature becomes relevant.

Our results demonstrate that the direction of a presented colored arrow could be recalled from memory despite being task irrelevant, though there was increased variability in precision. Furthermore, there was a measurable cost for adding direction to the memory set, even though only a single object was being stored. Thus, our data demonstrate a tradeoff between the quality of a memory and the number of features that need to be encoded even at a set size of 1 (see also Palmer, Boston, & Moore, 2015).

Our conclusion regarding the cost of adding features to memory appears to be inconsistent with some other results in the literature. For example, Olson and Jiang (2002) found in a change detection task that integrating features from different visual domains (e.g., size and orientation) does not produce a cost in memory performance. One possibility for the difference is that Olson and Jiang used change detection with categorically distinct stimuli, which may not have had the resolution to distinguish different levels of memory quality. Similarly, the results we found appear to be inconsistent with the results from Marshall and Bays (2013), who found equivalent memory precision for conditions in which one or two features of an oriented color bar were task relevant. However, there are two key differences between our design and theirs. First, four objects were presented in their task compared with one in ours; this may have caused responses to be too

noisy to accurately detect a moderate cost. Second, in their single-feature condition, participants had to simultaneously remember the color of one set of objects and the orientation of the other set of objects. The requirement of remembering one feature from one object and a different feature from a different object may not have been achievable by the participants, or it may have been less effortful for participants to encode both features for all four objects than to try to allocate memory resources differently between the two sets of objects.

Another recent finding from our lab revealed very poor accuracy in reporting task-relevant features of an attended object that participants did not expect to report (Chen & Wyble, 2015). That work used a four-alternative forced choice and thus could not measure memory precision. However, participants were often nearly at chance in answering the surprise questions, which suggests that almost all of the subjects were guessing. Furthermore, the task used by Chen and Wyble had four stimuli instead of one. The distinction between that study and the present one suggests that when participants are presented with multiple objects, other features may consume additional memory resources, thus further reducing the resources allocated to irrelevant features. For example, the number of objects and their spatial distribution may also consume resources even though that information is irrelevant. In support of this idea, it has been shown that individuals store ensemble statistics when shown a group of objects (Brady & Alvarez, 2011). It is important to consider that a great deal of irrelevant information is present in even the simplest visual display. Future work will need to explore how task demands affect the encoding of these various forms of memory that typically are not measured.

Our results provide support for our coarse coding hypothesis, which predicted that participants would have some memory for the irrelevant feature, although memory for this feature would be impoverished relative to memory for a relevant feature. This finding argues against hypotheses in which participants store memory for features of an object in an all-or-none fashion (Luck & Vogel, 1997; Serences et al., 2009) and hypotheses that only relevant features are encoded (Awh et al., 2006; Chun, 2011). Our finding is corroborated by monkey neurophysiological research in which attention was found to be a modulator of the amplitude of a neuron's tuning curve (McAdams & Maunsell, 1999), which would alter the relative strength of different features in memory according to their relevance. However, contrary to the simplest coarse coding hypothesis (C in Figure 1), participants were variable in the precision of their memories for the irrelevant feature. It is worth noting that a limitation of the design presented here is that participants were pretrained on

how to respond to the initially task-irrelevant feature. However, the pretraining is essential to ensure that the coarse memory retrieval in the surprise trial was not due to participants' confusion about how to report direction. Yet, despite this training, participants responded with degraded precision when direction was thought to be irrelevant.

Our findings provide new constraints on how we should theorize about capacity limits of visual working memory. First, we demonstrated that different features of an attended object can be represented with variable levels of precision according to task demands, which builds on current theories that single features of different objects are encoded with variable levels of precision (van den Berg, Shin, Chou, George, & Ma, 2012). Additionally, the variability of precision for a given feature seems to be related to how that feature is encoded into long-term memory (Fan & Turk-Browne, 2013). This finding also extends previous research that has shown that a cued object can be retrieved with greater precision than an uncued object (Bays, 2014; Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011) to suggest that such cueing is similarly effective for features within an object. Our findings were obtained even though the information to be stored concerns only a single object, which is far below typical estimates of working memory capacity.

Our results also show that even distinct feature dimensions (i.e., color and direction) interact at the level of memory representations, as demonstrated by the reduction in memory precision for color once direction became relevant to the task. Whether this interaction occurs because of shared neural resources for memory operations as predicted by the binding pool (Swan & Wyble, 2014) or because features are sampled less often when there are two relevant features compared with one (Vul & Rich, 2010) currently cannot be determined from this data set alone.

Furthermore, our result showing that participants variably encode task-irrelevant information may relate to the findings of Vogel, McCollough, and Machizawa (2005), who found that higher working memory capacity is correlated with differences in tendency to filter out irrelevant objects. We demonstrate here that variability in the tendency to filter out irrelevant information also exists at the level of features within a single object.

Conclusions

The results of the present experiment demonstrate that task-irrelevant features of an attended object are not entirely disregarded, nor are they represented at the same level as relevant features. Instead, participants are

able to report the irrelevant feature in a surprise test, but with greatly varying levels of precision. When the task-irrelevant feature became relevant following the surprise test, its precision improved, but at the expense of another task-relevant feature. These data challenge strong object-based encoding theories, which do not allow for variation in the storage of different features within a stored object representation.

Keywords: visual working memory, delayed estimation, visual attention

Acknowledgments

The authors thank the undergraduate students in the lab for assisting with the experiment as well as Ronald van den Berg and an anonymous reviewer for helpful comments. All authors contributed to the analysis of the data and to the writing and editing of the article. GS designed the experiment and the visualizations. This work was supported by NSF Grant BCS-1331073 awarded to BW. Data from this experiment can be found at <https://scholarsphere.psu.edu/collections/qv33rw66f>.

Commercial relationships: none.

Corresponding author: Garrett Swan.

Email: gsp.swan@gmail.com.

Address: Department of Psychology, Pennsylvania State University, University Park, PA, USA.

Footnotes

¹ But note that, in general, a characterization of visual working memory capacity strictly in terms of number of objects is incomplete (e.g., Alvarez & Cavanagh, 2004; Hardman & Cowan, 2014; Ma, Husain, & Bays, 2014).

² All 175 participants were used to minimize the bias induced by the subject exclusion process.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.
- 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.

- Asplund, C. L., Todd, J. J., Snyder, A. P., Gilbert, C. M., & Marois, R. (2010). Surprise-induced blindness: A stimulus-driven attentional limit to conscious perception. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1372–1381.
- Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*, 201–208.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839.
- Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *The Journal of Neuroscience*, *34*, 3632–3645.
- Bays, P. M., Catalao, R. F., & 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, doi:10.1167/9.10.7. [PubMed] [Article]
- Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, *11*(10):6, 1–15, doi:10.1167/11.10.6. [PubMed] [Article]
- Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*, 1622–1631.
- Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, *22*, 384–392.
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5):4, 1–34, doi:10.1167/11.5.4. [PubMed] [Article]
- Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science*, *24*, 981–990.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Chen, H., & Wyble, B. (2015). The location but not the attributes of visual cues are automatically encoded into working memory. *Vision Research*, *107*, 76–85.
- Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, *49*, 1407–1409.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Science*, *24*, 87–114.
- Eitam, B., Shoval, R., & Yeshurun, Y. (2015). Seeing without knowing: Task relevance dissociates between visual awareness and recognition. *Annals of the New York Academy of Sciences*, *1339*, 125–137.
- Eitam, B., Yeshurun, Y., & Hassan, K. (2013). Blinded by irrelevance: Pure irrelevance induced “blindness.” *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 611–615.
- Fan, J. E., & Turk-Browne, N. B. (2013). Internal attention to features in visual short-term memory guides object learning. *Cognition*, *129*, 292–308.
- Fisher, N. I. (1995). *Statistical analysis of circular data*. Cambridge, United Kingdom: 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, 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, doi:10.1167/10.12.27. [PubMed] [Article]
- Gao, T., Gao, Z., Li, J., Sun, Z., & Shen, M. (2011). The perceptual root of object-based storage: An interactive model of perception and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1803–1823.
- Gao, Z., Li, J., Liang, J., Chen, H., Yin, J., & Shen, M. (2009). Storing fine detailed information in visual working memory—Evidence from event-related potentials. *Journal of Vision*, *9*(7):17, 1–12, doi:10.1167/9.7.17. [PubMed] [Article]
- Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 325–347.
- Horstmann, G. (2005). Attentional capture by an unaccounted color singleton depends on expectation discrepancy. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1039–1060.
- Horstmann, G. (2006). The time course of intended and unintended allocation of attention. *Psychological Research*, *70*, 13–25.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.
- Keshvari, S., van den Berg, R., & Ma, W. J. (2013). No

- evidence for an item limit in change detection. *PLoS Computational Biology*, *9*(2), e1002927.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*, 391–400.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*, 347–356.
- Marshall, L., & Bays, P. M. (2013). Obligatory encoding of task-irrelevant features depletes working memory resources. *Journal of Vision*, *13*(2):21, 1–13, doi:10.1167/13.2.21. [PubMed] [Article]
- McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *The Journal of Neuroscience*, *19*, 431–441.
- Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, *41*, 1212–1227.
- 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., Boston, B., & Moore, C. M. (2015). Limited capacity for memory tasks with multiple features within a single object. *Attention, Perception, & Psychophysics*, *77*, 1488–1499.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 261–282.
- Rock, I., Linnett, C. M., Grant, P., & Mack, A. (1992). Perception without attention: Results of a new method. *Cognitive Psychology*, *24*, 502–534.
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*, 207–214.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentive blindness for dynamic events. *Perception*, *28*, 1059–1074.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.
- Suchow, J. W., Brady, T. F., Fougny, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10):9, 1–8, doi:10.1167/13.10.9. [PubMed] [Article]
- Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, & Psychophysics*, *76*, 2136–2157.
- Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision*, *3*(1):9, 86–94, doi:10.1167/3.1.9. [PubMed] [Article]
- 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.
- 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.
- Vul, E., & Rich, A. N. (2010). Independent sampling of features enables conscious perception of bound objects. *Psychological Science*, *21*, 1168–1175.
- 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, 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 and Review*, *15*, 223–229.
- Xu, Y. (2010). The neural fate of task-irrelevant features in object-based processing. *The Journal of Neuroscience*, *30*, 14020–14028.
- Yin, J., Zhou, J., Xu, H., Liang, J., Gao, Z., & Shen, M. (2012). Does high memory load kick task-irrelevant information out of visual working memory? *Psychonomic Bulletin and Review*, *19*, 218–224.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 233–235.