Location Has a Privilege, but It Is Limited: Evidence from Probing Task Irrelevant Location

Joyce Tam[1] and Brad Wyble[1]

[1]Department of Psychology, Pennsylvania State University, University Park, PA 16801.

Author Note

This study is supported by funding from NSF grant 1734220 awarded to author B. W. Stimuli, experiment script and data are available at https://osf.io/8k79z/

Correspondence concerning this article should be addressed to Joyce Tam, Department of Psychology, Pennsylvania State University, University Park, PA 16801. Email: joyce.m.y.tam@gmail.com.

We have no known conflict of interest to disclose.

Abstract

We investigated the extents of automaticity in location and orientation encoding in visual working memory (VWM) by manipulating their task relevance and assessing the amount of resource recruited by their encoding. Across five experiments, participants were surprised with a location report trial (Experiment 1A, 2A, and 3) or an orientation report trial (Experiment 2A and 2B) at a point when only the item's color had been task relevant. This was followed by control trials to assess the memory quality of color when location or orientation had become task relevant. We found the surprise trial performance to be significantly worse than the first control trial for both location and orientation, although to a greater extent for orientation for which there was virtually no measurable information from the subjects' reports. This was the case even when encoding was the only incidental memory process before the control trials (Experiment 2A and 2B), and the surprise memory costs cannot be attributed to the unexpectedness inherent to the surprise question (Experiment 3). The control trials revealed a consistent reduction of color memory only in the orientation experiments. These results suggest that although location encoding is more automatic than orientation, neither is encoded in a fully automatic manner. Our results show that incidentally encoded location is only coarse-grained, constraining the spatial precision of space-based indexing systems.

Keywords: visual working memory, automaticity, delayed estimation

Location Has a Privilege, but It Is Limited: Evidence from Probing Task Irrelevant Location

Visual working memory (VWM) provides a means to rapidly store readily accessible, task relevant visual information. However, to understand the *automaticity* of information encoding, it is essential to study the memory of visual information when it is task irrelevant. To illustrate, the classic dual-mode model (Schneider & Shiffrin, 1977) postulates an automatic process to be a sequence of subprocesses that would "(nearly) always become active in response to a particular input configuration" (p. 2). This contrasts with controlled processes, where the sequence of subprocesses is activated through intentions and thus requires control resources. Accordingly, automatically encoded information should survive at least two tests: First, its task relevance should have no bearing on how well it is encoded, as the sequence of subprocesses underlying its encoding should always activate regardless of the task requirements, such that all available information would be encoded incidentally. The second test is that encoding this information should incur no evident cost to other cognitive operations (and vice versa), as an automatic process will recruit no control resources (see also Moors & De Houwer, 2006).

An example of automaticity in memory encoding is demonstrated in the contextual cueing effect (Chun & Jiang 1998), where spatial configuration is encoded even when it is not explicitly task relevant. After consistent mappings between the location of a search target and a specific embedding spatial configuration, search is sped up when the target is presented in the associated configuration, even if no explicit memory of that configuration is formed (Chun & Jiang, 1998). This finding suggests that spatial contexts can be implicitly acquired over repeated exposure and are represented along with the target (Chun, 2000; Chun & Jiang, 1998). In VWM paradigms, however, target locations and other attributes often vary unpredictably from trial to trial. Thus, for models organizing working memory representations on spatial reference frames (e.g., Bays, 2014), location encoding must not only occur incidentally, but such incidental encoding must also be flexible and rapid (on the scale of individual trials). Therefore, the encoding of location regardless of its task relevance is a necessary assumption behind these models, whereas resource-free location encoding would be a helpful support to their cases. Similarly, rapid and incidental encoding of all other task irrelevant attributes of an attended item would be necessary if VWM representations have to be object-based (e.g., Luck & Vogel, 1997). In the following pages we will review past studies discussing the automaticity of the trial-by-trial encoding of location and other features (e.g., color, shape, spatial frequency, orientation) with the lens of the two tests of automaticity from the dual-mode framework (i.e., whether the information is encoded when task irrelevant, and whether such encoding requires a substantial amount of resource). For convenience, in the rest of the article, we will refer to features other than location and the updating of location (i.e., movement) simply as “nonspatial” features.

The effect of task relevance on VWM representations has been studied with tests of memory that are explicit (i.e., directly probing a task irrelevant feature) or implicit (i.e., measuring the effect of a task irrelevant feature through its influence on the ongoing task). In studies using an explicit task, task relevance seems to have minimal effect on location memory but strongly modulates the memory of nonspatial features. Participants who were not instructed to remember location did not remember it any worse than those who knew location was task relevant (Mandler et al., 1977; McCormack, 1982; Pezdek & Evans, 1979; Schulman, 1973; Smith & Milner, 1989; Zechmeister et al., 1975), nor were participants any worse at reporting location in a surprise trial compared to the next post-surprise trial (Chen & Wyble, 2015a, 2016, 2018). However, unexpectedly probing a task irrelevant nonspatial feature resulted in very low accuracy (Chen & Wyble, 2015a, 2015b, 2016, 2018; Eitam et al., 2013; Shin & Ma, 2016; Wyble et al., 2019). Non-dichotomous manipulation of task relevance through test probability showed similar results, as the probability of being tested modulates memory for shape and color (Pilling & Gellatly, 2013; Van Lamsweerde & Beck, 2011) but not location (Van Lamsweerde & Beck, 2011).

In contrast, studies using implicit tests of memory provided evidence that task irrelevant features, be they spatial or nonspatial, can be remembered to some extent. For example, an irrelevant feature of an item currently being maintained in VWM can attract attention and influence a concurrent visual search task – and such evidence has been found for location (Hollingworth & Bahle, 2020; De Vries et al., 2018; but see Sala & Courtney, 2009; Van Moorselaar et al., 2014) as well as nonspatial features (Foerster & Schneider, 2018, 2020; Gao et al., 2016; Hollingworth & Luck, 2009; Hollingworth et al., 2013a, 2013b; Soto & Humphreys, 2009; Thayer et al., 2021; but see Olivers et al., 2006; Sala & Courtney, 2009). A related line of studies found that memory of the relevant feature can be impaired if a mismatch exists between the study and the test display along an irrelevant feature dimension – both when the irrelevant mismatch occurs spatially (Gu et al., 2020; Jiang et al., 2000; Rajsic & Wilson, 2014; Vidal et al., 2005; but see Woodman et al., 2012) or along a nonspatial feature dimension (Ecker et al., 2013; Gao et al., 2010; Hyun et al., 2009; Yin et al., 2012; Yin et al., 2011; but see Shin & Ma, 2017). In addition, some studies showed that irrelevant location information influences the ongoing task more readily, consistent with superior memory for location in the explicit studies. Directly comparing the amount of implicit influence from irrelevant location and irrelevant nonspatial features, Logie et al. (2011) showed that the effect of irrelevant mismatching of location, color, and shape all dissipated at longer study-test intervals, but the influence from location mismatch was the most persistent (see also Bocincova et al., 2017[1]). In addition, feature misbinding errors were elevated when items incidentally shared the same location but not when they incidentally shared the same color (Pertzov & Husain, 2014; Schneegans et al., 2021).

Turning now to resource recruitment as a means to assess automaticity, remembering an additional feature in VWM has been shown to take up additional cognitive resources, no matter whether the feature is spatial or nonspatial. Memory costs have been observed when the number of to-be-encoded features increased, both in the case where the tested features were nonspatial (Cowan et al., 2013; Fougnie et al., 2010; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013; Swan et al., 2016; but see Luck & Vogel, 1997; Shin & Ma, 2017; Vogel et al., 2001), or spatial (Palmer et al., 2015). Furthermore, dual-task costs have been observed when participants were maintaining memory of nonspatial features (Cowan & Morey, 2007; Ricker et al., 2010) or spatial features (Oh & Kim, 2004; Woodman & Luck, 2004). Consistently, the amplitude of the contralateral delay activity (CDA), which increases with working memory load (Vogel & Machizawa, 2004), was found to increase to a similar extent when color or location was the relevant feature (Hakim et al., 2019). Still, the CDA has been shown to correlate with the number of to-be-remembered items but dissociate with the feature load (Quak et al., 2018), and thus the CDA increase observed in Hakim et al. (2019) may reflect a feature-nonspecific recruitment of resources modulated by the memory set size.

Summary and Current Study

---

[1] Boncincova et al. (2017) replicated this data pattern reported by Logiet et al. (2011) but the interpretations they provided were quite different. Logie et al. (2011) posited that all features were initially bound in perceptual/ sensory memory, but irrelevant features were gradually inhibited during the formation of VWM. However, Bocincova et al. (2017) showed that the dissipation of the irrelevant change effect can be entirely accounted for by the gradual decay of the sensory memory trace, as the irrelevant change effect was eliminated across all study-test intervals when the sensory trace of the memory array was eliminated via masking.

From these studies, we learn that spatial information is typically easy to report whether it is task relevant or not, but the case for nonspatial information is more nuanced: When nonspatial information is task irrelevant, there is poor accuracy in explicit memory tests, but implicit measures suggest that at least some of this information is maintained. On the other hand, a requirement to remember spatial or nonspatial information were both shown to recruit additional resources, resulting in a cost to memory performance. Therefore, while both types of information show automaticity in certain ways, location memory has more substantial, though not unequivocal, evidence of being automatic.

It is also important to note that different sub-processes of memory, such as encoding, maintenance and retrieval may differ in automaticity. If evidence for full automaticity were present for both tests of automaticity described above (that memory was unaffected by task relevance and did not recruit additional resources) then we could infer that all memory sub-processes were fully automatic. However, a failure to observe full automaticity could be linked to any of these sub-processes, requiring further study.

Thus, the purpose of the current study is twofold: (1) to test whether the incomplete automaticity in location memory as suggested by the resource recruitment evidence can be demonstrated as a task relevance memory effect if a highly sensitive continuous estimation test is used, and (2) to specifically investigate the automaticity of the *encoding* process for both spatial and nonspatial information by minimizing the influence of post-encoding processes. In addition to referencing the two tests of automaticity from the dual-mode framework (Schneider & Shiffrin, 1977), we also adopted a graded, nonbinary view of automaticity (Hasher & Zacks, 1979; Logan, 1985), such that the degree of which processes are gated by task relevance and the amount of resource they recruit can vary along a continuum.

With these in mind, we chose to use a surprise test paradigm coupled with a delayed estimation task as in Swan et al. (2016). In their study, participants were instructed to remember only the color of a stimulus and perform the color task for 25 trials, before being unexpectedly probed to report an irrelevant feature, orientation, on the 26th trial. After the surprise trial, either the always-relevant feature (color) and the additionally required, *critical feature* (orientation), were tested randomly on each trial. This paradigm is instructive for the current purpose because it encompasses the two tests of automaticity: The comparison between the surprise trial and the first post-surprise trial for the critical feature tests whether task relevance affects the memory of that feature, and the comparison of color reports in the pre-surprise and the post-surprise trials tests whether the additional requirement to remember the critical feature induces a cost to the memory of the always-relevant feature. This paradigm also allows us to change the encoding status of features without modifying the complexity of the display.

In the current study, we assessed and compared the automaticity of the encoding of location and orientation across five experiments. Except for Experiment 3, the always-relevant feature was color. The critical feature was location in Experiment 1A and 2A, and it was orientation in Experiment 1B and 2B. Location was always the tested feature in Experiment 3. Experiment 1A and 1B showed that both orientation and location was reported less accurately on the surprise trial compared to the first post-surprise trial, although this difference was larger for orientation. In addition, color memory grew worse in the post-surprise trials in both Experiment 1A and 1B. The incomplete automaticity cannot be fully attributed to post-encoding processes such as active inhibition of irrelevant features (Experiment 2A and 2B, see Logie et al., 2011) or interference from the unexpectedness inherent to the surprise trial (Experiment 3, see Swan et al., 2017). Taken together, our results suggest that location encoding is more automatic than orientation encoding, but neither are fully automatic.

Experiment 1A and 1B

Methods

The experiment scripts, data, and replication registration are available at https://osf.io/8k79z/. While the experimental design complied with the pre-registerded plan, some deviations exist in the analysis methods. This includes the use of the two-component mixture model (Zhang & Luck, 2008) instead of the target confusability competition model (Schurgin et al., 2020) to fit the error data from the continuous estimation task, the additional use of the replication Bayes Factors (Ly et al., 2019; Verhagen & Wagenmakers, 2014), and the omission of permutation analyses in comparing single-trial accuracies. These analysis methods will be detailed and justified below. Additional description of these deviations, alongside the descriptions of the initial experiments which Experiment 1A and 1B replicate, can also be found in the supplemental materials. Note also that Experiment 2A, 2B, and 3 are *not* pre-registered nor are they replication experiments.

The experiments were programmed in JavaScript with the PsychoJS library and hosted with Pavlovia (Peirce et al., 2019). Analyses were aided with R v4.0.2 (R Core Team, 2020), the MemToolbox (Suchow et al., 2013) and JASP v0.14 (JASP Team, 2020). The experiments were approved by the institutional review board at the Pennsylvania State University.

Participants

Experiment 1A had a sample of 77 (M = 18.59 years, SD = 0.83 years) and experiment 1B had a sample of 80 (M = 18.93 years, SD = 1.71 years). These sample sizes were set to approximate the sample size of a separate set of initial experiments, the design of which Experiment 1A and 1B each replicate (see supplemental materials). 100 participants were recruited for experiment 1A, with 17 being excluded for poor performance in the response training task and 6 excluded for poor performance in the main task; 135 participants were recruited for experiment 1B with 47 excluded for the response-training task and 8 for the main task (see Design for task descriptions and exclusion criteria). All participants were recruited from the Penn State University subject pool in exchange for course credits.

Design

The major variables of interest were assessed in the main task, which consisted of a surprise trial paradigm coupled with a delayed estimation task. To ensure that participants understood how all features can be reported, the main task was preceded with a response-training task in each of the experiments. Note also that the experiments were completed online, so that stimuli sizes and distances are reported in pixels instead of visual angles.

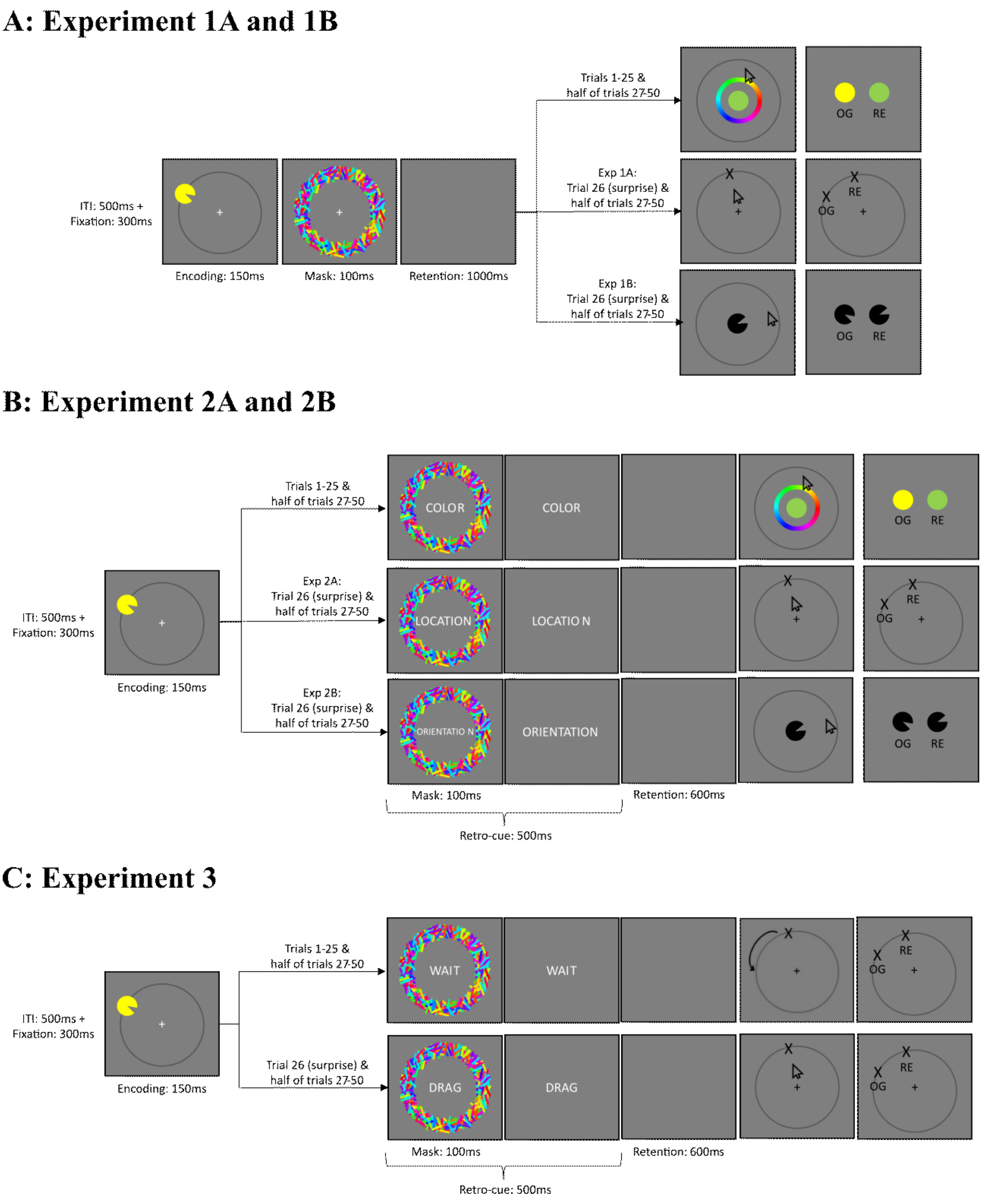
Response-Training Task: Perceptual Matching. The response-training task was identical in experiment 1A and 1B. It consisted of blocks of perceptual matching trials of color, location, and orientation (10 trials each). On color trials, a uniformly colored circle (radius = 50 px) is presented on the left half of the screen. Another circle of the same size was placed on the right half and was surrounded by a color wheel (radius = 125 px). The color of the right circle varied with the cursor's position relative to the color wheel with colors spanning through the entire hue dimension in the HSV color space (with full saturation and lightness value). Participants were to use their cursor to adjust the color of the right circle such that the colors of the circles matched as well as could be achieved. The location and orientation matching trials were similar. On location trials, a white circle (radius = 50 px) marks a position on a centered circular ring (radius = 300 px), and participants used their cursor to adjust the position of an “X” (height = 40 px) on the same circular ring to match its position to the circle’s. On orientation trials, a “pacman” in black (i.e., a circle with a wedge removed, radius = 50 px, sector angle = 20°) was presented on the left. Participants were to use their cursor to adjust the orientation of another black pacman on the right so that the two pacmen have the same orientation. The feature value of the match target for each feature would be randomly chosen from a pre-generated sequence of values equally distributed on the circular scale. Unlimited time and chances were provided until participants were satisfied with the match and pressed the spacebar to proceed. The next trial would begin after 500ms. The trials were blocked by the feature type and block order was randomized across participants.

Errors were calculated by subtracting the match target’s value from the reported value. Participants with absolute errors larger than 10° in any of the last three trials requiring perceptual matching of the critical feature (i.e., location for participants in experiment 1A and orientation for participants in experiment 1B) were excluded from further analyses. This strict criterion follows the decision by Swan et al. (2016) and aims to minimize the opportunity that participants could not report the critical feature on the surprise trial in the subsequent delayed estimation task because they were unable to use the reporting method for any reason.

Main Task: Delayed Estimation. The delayed estimation task was labelled as “Experiment 2” to help participants establish that they should not carry expectations forward about the requirement to report orientation or location into the main task. The trial sequence is depicted in Figure 1A. The critical feature in the main task was location for experiment 1A and orientation for experiment 1B but the task was otherwise identical. Each trial started with a white fixation cross (height = 20 px) at the screen center for 300ms, before the to-be-remembered item was presented for 150ms. The item varied randomly and uniformly in color, location, and orientation: It had a pacman-like shape, with a radius of 50 px and a sector spanning 20°. While the sector remained uncolored, the remaining portion had a random color from the color wheel. The orientation of the sector would vary randomly from 1 to 360°. In addition, it sat on a random location on a centered circular outline (radius = 300 px). The memory item was then masked for 100ms. After a 1000-ms blank delay interval, the fixation and the outline circle reappeared, and participants reported the memorized value of the probed feature attribute.

Participants were instructed to only remember the item's color and they would first complete 25 color pre-surprise trials. On the 26th trial (the surprise trial), the critical feature was unexpectedly probed using the same procedure as in the response-training trials that they had previously completed, except that they were now responding based on memory. Then, for the 27th to the 50th trial (i.e., the post-surprise trials), color or the critical feature was randomly probed with equal chance. The test screen that appeared after the 1000-ms blank delay varied according to the probed feature. On color probe trials, a circle (radius = 50 px) would be presented at the screen center surrounded by a color wheel (radius = 125 px). The circle's color was adjustable with the cursor's position. On location trials (only present in experiment 1A), a black "X" (height = 40 px) was presented on the outline circle, and the location of the "X" was adjustable with the cursor. On orientation trials (only present in experiment 1B), a black pacman with cursor-adjustable orientation was presented at the screen center. Participants were asked to adjust the probed feature to match the remembered feature as well as possible. A feedback screen with both the correct value and the reported value would then be presented until participants pressed the spacebar again. The next trial would begin 500ms afterwards. On the surprise trial, a text string was presented 160 px below the screen center concurrent to the test display onset and read “This is a surprise test. What was the [critical feature] of the stimulus?”, with the text for the critical feature being “location” in experiment 1A and “orientation” in experiment 1B.

Similar to the perceptual matching task, errors were calculated by subtracting the memory item’s value from the reported value. The data from each experiment was grouped by the tested feature attribute (color or critical feature) and the trial type (pre-surprise, surprise, or post-surprise).


A: Experiment 1A and 1B

B: Experiment 2A and 2B

C: Experiment 3

Figure 1. Trial Sequence of the delayed estimation task in different experiments. In Experiment 1A and 1B, location or orientation was unexpectedly probed when color was considered the only relevant feature. Experiment 2A and 2B followed this task structure, but the to-be-probed feature was always validly cued immediately after encoding. Experiment 3 always probed location, but the response method unexpectedly changed on the surprise trial.

Analyses

Model Fitting. For the main task, in addition to analyses using the mean absolute errors, maximum likelihood estimates (MLE) were obtained by fitting the error values to the two-component mixture model, which assumes a uniform distribution reflecting the probability of guessing (P_U) and a von Mises distribution reflecting the probability that the target was indeed memorized, and its concentration parameter (s) reflects memory precision (Zhang & Luck, 2008)[2]. For multi-trial data, i.e., data from pre-surprise trials and post-surprise trials, models were fitted per participant per condition. For single-trial data, i.e., data from the surprise trial and the first post-surprise trial, models were fitted per condition after collapsing across participants[3][4]. For each pair of fitted parameters, we used functions from the MemToolbox (Suchow et al., 2013) to obtain Markov Chain Monte Carlo samples from the posterior distribution of s and P_U respectively. The posterior samples are then used to form a 95% credible interval for each estimate.

---

[2] In the Results section, these parameters are interpreted according to the assumption of the model proposed by Zhang & Luck (2008). We note that the two-component assumption may not be shared by all researchers. For example, Schurgin et al. (2020) recently proposed a single-component model (the target confusability competition model) and suggested that the presence of guessing need not be assumed. When feature values on the report wheel are scaled with respect to their perceived similarity with the target (instead of their absolute similarity), one free parameter, the sensitivity index *d'* suffices. Data analyses with the target confusability competition model, where applicable, are available in the supplemental materials.

[3] The data points in each condition were also fitted in the same manner to single-component models that either only allows s to vary or only allows P_U to vary. With the Akaike Information Criterion, we found the two-component model to be the best-fit model in 13 out of the 18 total conditions throughout the paper, with an average advantage of 18.29 units over the s-only model and 139.45 units over the P_U only model. In the remaining 5 conditions, the s-only model has the best fit, but only with a small advantage over the two-component model (average advantage 1.26 units). In other words, the two-component model is either the best-fit model or has a comparable fit as the best-fit model, and thus, we reported parameters fitted with the two-component model throughout the paper. See supplemental materials for details of this model comparison process.

[4] The number of observations to generate model estimates for each single-trial condition is equal to the sample size. We ran simulation analyses to show that the sample sizes in the experiments render adequate numbers of observations to provide reliable model estimates. See supplemental materials for details.

In addition to the exclusion criteria based on the performance in the response-training task, participants were excluded if their performance in the delayed estimation task was poor for color or the critical feature in the pre-surprise or post-surprise trials (mean P_U exceeding 0.4 or mean s exceeding 50°, following the criteria used by Swan et al., 2016).

Bayesian Analyses. As Experiment 1A and 1B are replication experiments with the same design as a separate set of initial experiments (see supplemental materials), we report the replication Bayes factor, which denotes the amount of evidence for or against the presence of an effect from the replication study when the prior distributions have been updated with the initial results (Verhagen & Wagenmakers, 2014). While direct alterations of the prior distribution can be performed, the replication Bayes factor is also simply the quotient when the combined Bayes factor (i.e., the Bayes factor computed from data from the initial and replication experiments combined) is divided by the initial Bayes factor, or:

$$\mathrm{BF}\left(d_{rep}|d_{ini}\right) = \mathrm{BF}\left(d_{ini}, d_{rep}\right) / \mathrm{BF}(d_{ini})$$

where d_{ini} and d_{rep} refer to the data from the initial study and the replication study respectively (Ly et al., 2019, see also Harms, 2019). For comparisons between two samples, Bayes factor denoting the amount of evidence in favor of the alternative hypothesis, BF_{10} is reported. For factorial analyses, we report BF_{incl} which denotes the amount of evidence for including the effect of interest in the model. The initial BFs and combined BFs can be found in the supplemental materials.

Critical Feature Memory. Memory of the critical feature on the surprise trial (SUR) was compared with the first post-surprise trial (FPS) to assess the effect of task relevance, with the mean absolute error compared by a paired Bayesian t-test and the model parameters compared through differences in their MLEs and overlaps between their 95% CIs. SUR and FPS yielding comparable performances would indicate that the encoding process is fully automatic, whereas finding memory on SUR being worse than FPS would indicate that one or more of the memory subprocesses have incomplete automaticity. In that case, further experiments would be needed to find out whether this memory cost is at least partly attributable to an encoding process that is not fully automatic. As we do not have the theoretical grounds to predict SUR to be better than FPS, all tests comparing FPS and SUR are one-tailed (H1: memory in FPS is better than memory in SUR). To account for the possibility that the SUR-FPS difference is exaggerated by participants deploying additional cognitive efforts on the first compared to subsequent post-surprise trials, comparison procedures for SUR and FPS were repeated to compare SUR with a random post-surprise trial (RPS) for experiments where a SUR to post-surprise improvement is predicted. Details of this analysis can be found in the supplemental materials.

To investigate any qualitative difference between SUR and FPS, we also compared the amount of categorical bias (Bae et al., 2015; Ester et al., 2020; Pratte et al., 2017) in these two trials. To this end, we computed the distance between the reported location or orientation value to its *nearest canonical value* (0°, 90°, 180°, or 270°) and used a two-tailed paired Bayesian t-test to determine whether this distance is different between conditions (a smaller distance from the nearest canonical value us taken as evidence for a stronger categorical bias). In addition, the empirical cumulative distribution function from the SUR error distribution was compared with a theoretical cumulative distribution function of a uniform distribution ranging from -180 to 180 with the Kolmogorov-Smirnov test to assess whether SUR memory is distinguishable from a noninformative, uniform distribution.

Color Memory. Color memory in the pre-surprise (Pre) and post-surprise trials (Post) were compared to assess whether the additional requirement to encode the critical feature recruited additional cognitive resources, resulting in a tradeoff against color memory. The logic is similar to the SUR-FPS comparison: Pre being indistinguishable from Post would indicate that the encoding for the critical feature is fully automatic, whereas Pre being worse than post would indicate that one or more of the memory subprocesses are not fully automatic, requiring further experimentations to narrow down the process(es) that contributed. The mean absolute error, MLE of s, and MLE of P_U from Pre and Post will be compared by paired Bayesian t-tests. Again, one-tailed tests were performed as we do not expect Post to be better than Pre in any scenarios (H_1: memory in Post is worse than memory in Pre).

Results

Effects of Task Relevance

Data from experiment 1A shows that location memory improved significantly from SUR to FPS (Figure 2A). The mean absolute error in SUR was 34.29° (SE = 0.53°), which decreased to 10.03° in FPS (SE = 0.75°), BF_{10} = 1.61E+04. The MLE of both parameters from the two-component mixture model reflected an improvement from SUR to FPS. The precision parameter s decreased from 20.00° (95% CI [15.37, 26.06]) to 11.81° (95% CI [10.01, 14.70]), suggesting an improvement in memory precision. The probability of guessing, P_U, also decreased from 0.24 (95% CI [0.13, 0.37]) to 0.02 (95% CI [0.00, 0.08]). The Kolmogorov-Smirnov test shows that location memory in SUR was significantly different from a uniform distribution, D = 0.31, $p < .001$. The evidence is ambiguous as of whether the amount of categorical bias changed between SUR and FPS, distance from the nearest canonical value in SUR: 34.75° (SE = 1.78°), in FPS: 34.09° (SE = 1.86°), BF_{10} = 0.75.

Experiment 1B shows that orientation memory also improved significantly from SUR to FPS (Figure 2B). The mean absolute error decreased from 80.35° (SE = 5.81°) to 28.79° (SE = 3.94°), BF_{10} = 1.45E+07. Model estimates reflect an improvement in memory precision, where s decreased from 69.27° (95% CI [65.28, 1094.79]) to 19.79° (95% CI [16.00, 27.30]), as well as an increased probability of memory being present, as P_U decreased from 0.73 (95% CI [0.01, 0.97]) to 0.20 (95% CI [0.10, 0.32]). However, the wide confidence intervals in parameter estimates for SUR suggest that the two-component mixture model does not describe the data in the surprise trial well. In fact, the Kolmogorov-Smirnov test shows that SUR was indistinguishable from pure guessing, D = 0.09, $p = .585$, This is also evident from visually inspecting Figure 2B. Again, the evidence is ambiguous on whether the amount of categorical bias changed between SUR and FPS, distance from the nearest canonical value in SUR: 28.20° (SE = 1.53°), in FPS: 28.01° (SE = 1.52°), BF_{10} = 0.44.

Furthermore, the improvement in memory from SUR to FPS was greater for orientation (Experiment 1B) than location (Experiment 1A), as a mixed Bayesian ANOVA suggests that the interaction effect between critical feature (between-subject: location or orientation) and trial type (within-subject: SUR or FPS) should be included in the model, BF_{incl} = 42.30. See Figure 3 (solid lines).

Effect of Resource Recruitment

In experiment 1A, color memory got worse from the Pre trials to the Post trials, providing additional evidence that location memory encoding is not fully automatic (Figure 4A). The mean absolute error for color increased from 13.59° (SE = 0.53°) to 15.87° (SE = 0.75°), BF_{10} = 99.05. While s showed a numerical increase from 14.96° (SE = 0.49°) to 16.34° (SE = 0.65°), the evidence was not strongly in favor or opposed to the presence of a decrease in memory precision, $BF_{10} = 0.82$. Still, P_U showed a very small but evident increase from 0.035 (SE = 0.009) to 0.049 (SE = 0.005), $BF_{10} = 8.13$, such that the probability of memory absent was slightly elevated despite being very low in both cases.

Experiment 1B shows that color memory was also worse in Post trials compared to Pre trials (Figure 4B), which is consistent with the earlier observation of the absence of automaticity in orientation memory. The mean absolute error for color increased from 12.71° (SE = 0.61°) to 16.10° (SE = 0.75°), $BF_{10} = 2271.45$. The parameter s also increased from 13.69° (SE = 0.40°) to 16.01° (SE = 0.66°), $BF_{10} = 84.80$, reflecting a decrease in memory precision. P_U also increased slightly and less evidently, from 0.036 (SE = 0.007) to 0.057 (SE = 0.010), $BF_{10} = 2.35$.

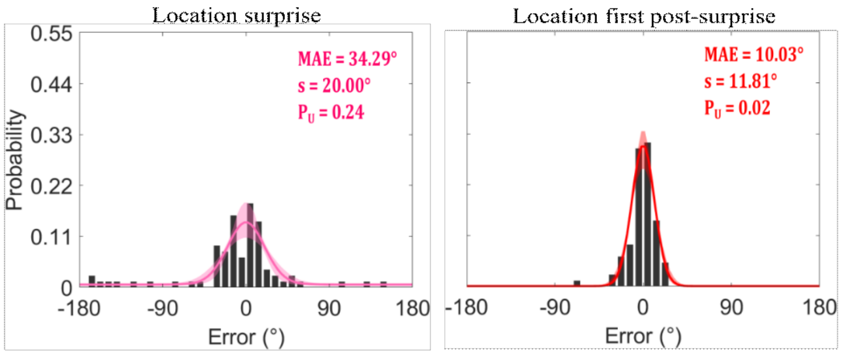
To compare Experiment 1A and 1B, we performed a mixed Bayesian ANOVA with critical feature (between-subject: location or orientation) and trial type (within-subject: Pre or Post) using each measurement (mean absolute error, s and P_U). The BF_{incl} for the interaction term between critical feature and trial type is taken as evidence to support that shifting memory of location and orientation from incidental to intentional caused different degrees of tradeoff with color memory. However, the evidence is mostly inconclusive (mean absolute error: $BF_{incl} = 1.05$, s: $BF_{incl} = 1.05$, P_U: $BF_{incl} = 0.53$).

Discussion

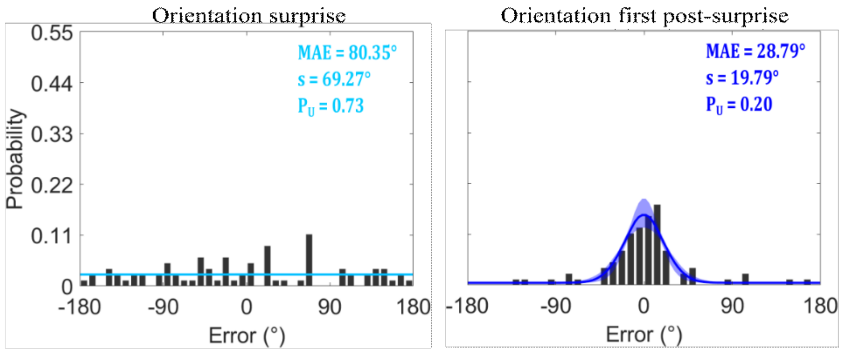
Data from Experiment 1A and 1B shows that neither location nor orientation memory could be remembered in a fully automatic manner, although location memory showed a higher degree of automaticity. Unknown is which memory subprocess(es) contributed to this lack of full automaticity. To illustrate, the observed task relevance effect could have been because less information about the critical feature was encoded when it was deemed task irrelevant (i.e., incomplete automaticity in the encoding process), but it could also have been because of active inhibition of the critical feature due to its task irrelevance (Logie et al., 2011), or exclusion from active maintenance in the focus of attention.(Cowan, 2001; Oberauer, 2013) when it was task irrelevant (i.e., incomplete automaticity in post-encoding processes). Of course, it could have been that both encoding and post-encoding processes contributed to this effect. Similarly, the resource recruitment cost might have been contributed by one or more of these processes, and they do not have to match those contributing to the task relevance effect. Moreover, the automaticity of these memory subprocesses may differ for location and orientation.

To isolate the automaticity of encoding to post-encoding processes, in Experiment 2A and 2B, the to-be-tested feature was indicated by an always-valid retro-cue (Park et al., 2017; Souza & Oberauer, 2016) presented immediately after the memory display. Therefore, on the surprise trial, the cue would make the critical feature task relevant immediately after the memory display rather than after the retention interval. Replicating the task relevance effect from the SUR-FPS comparison in Experiments 1 would strongly suggest that the effect was not caused by effortful post-encoding processes such as active inhibition or the lack of active maintenance of task irrelevant information. Also, on color trials, the cue would indicate color as the relevant feature immediately after the memory display, eliminating any motivation to maintain the critical feature further, making the encoding of the critical feature the only added requirement post-surprise. Therefore, observing a resource recruitment cost from the color Pre-Post comparison would strongly indicate that the encoding process of the critical feature was not fully automatic.

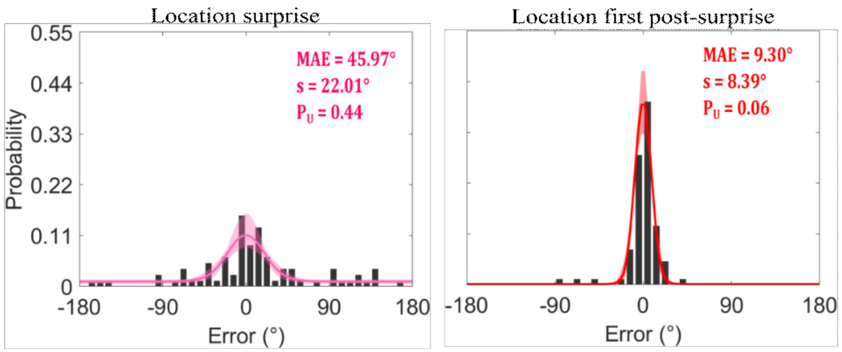
A: Experiment 1A



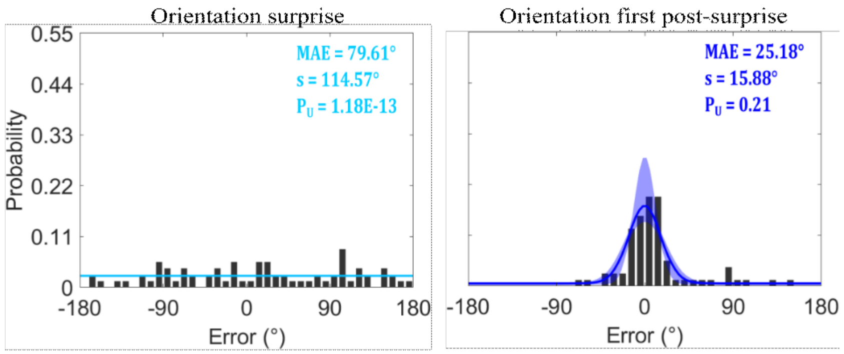
B: Experiment 1B



C: Experiment 2A



D: Experiment 2B



E: Experiment 3

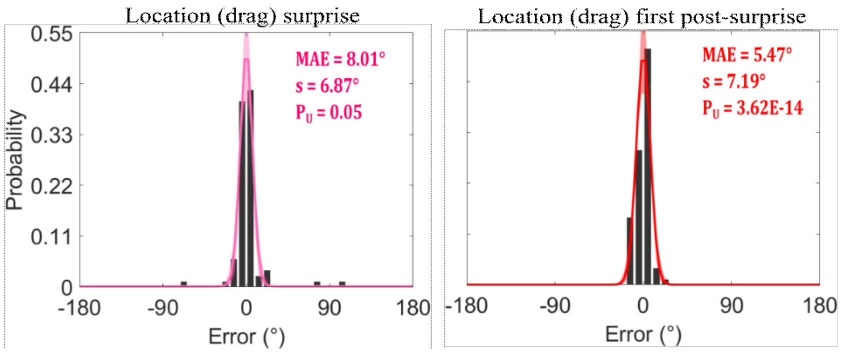


Figure 2. The effect of task relevance on location encoding and orientation encoding across experiments. Location memory demonstrated a superior (to orientation), but incomplete degree of automaticity (**A** and **B**). Their incomplete automaticity can be attributable to the encoding process (**C** and **D**), but not interference from the unexpectedness of the surprise trial (**E**). MAE stands for mean absolute errors. s and PU are MLE of the two-component mixture model. Shaded areas are 95% credible intervals.

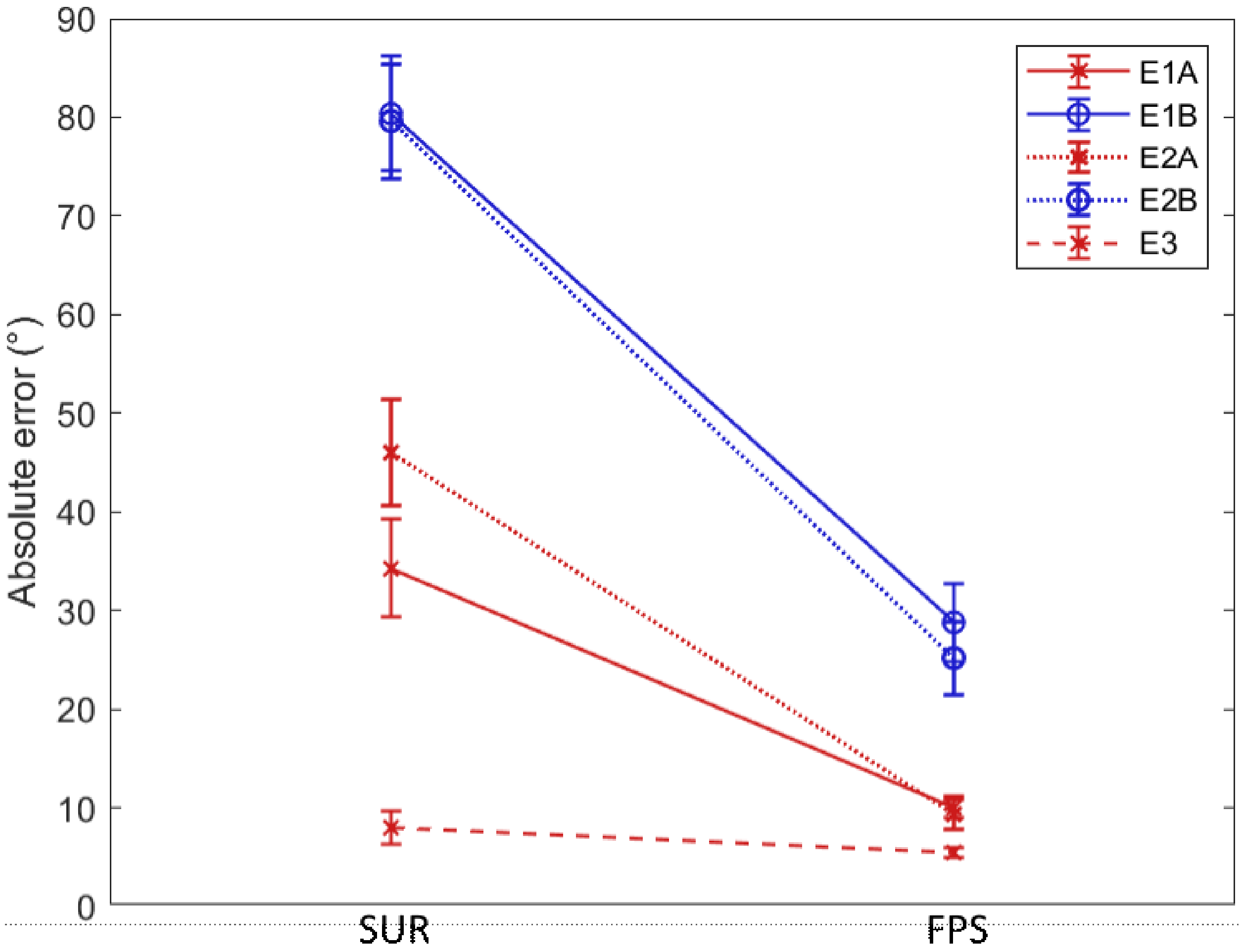
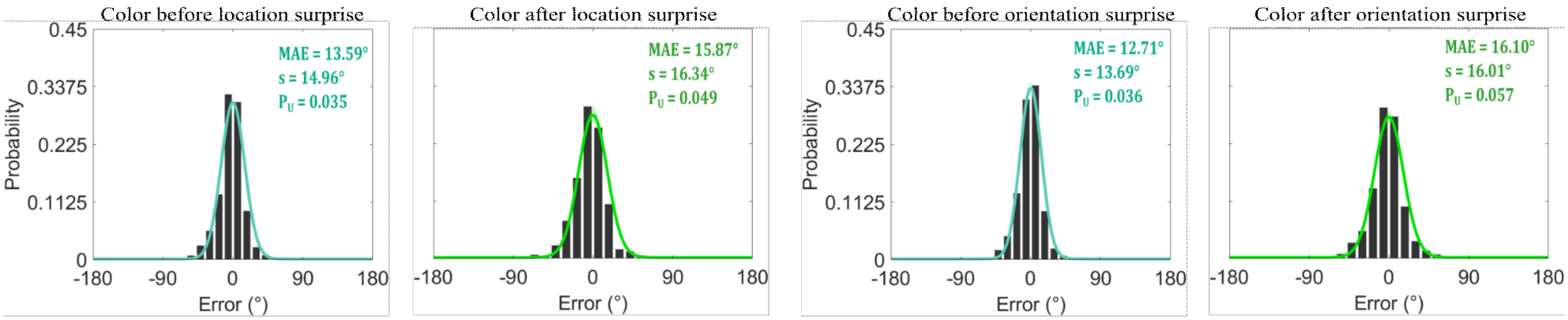
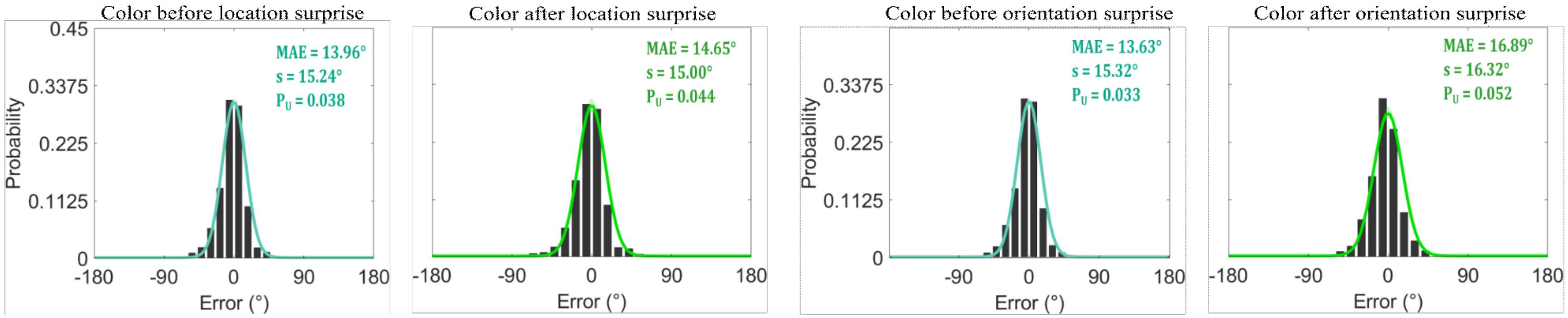


Figure 3. The differential effect of task relevance on location and orientation encoding across experiments in terms of absolute error (location data from Experiment 1A, 2A and 3 in red, and orientation data from Experiment 1B and 2B in blue). SUR stands for surprise and FPS stands for first post-surprise. Error bars are mean standard errors.



A: Experiment 1A

B: Experiment 1B



C: Experiment 2A

D: Experiment 2B

Figure 4. Change in color memory when location or orientation was the critical feature across experiments. While the additional requirement to remember location and orientation both led to a tradeoff in color memory (**A** and **B**), this cost became less evident when only the encoding of location was additionally required (**C**). In contrast, the cost to color remained when only the encoding of orientation was additionally required (**D**). MAE stands for mean absolute errors. s and PU are MLE of the two-component mixture model. Shaded areas are 95% credible intervals.

Experiment 2A and 2B

Methods

Unless specified below, the methods for Experiment 2A and 2B were identical to those for Experiment 1A and 1B. Experiment 2A had a sample of 78 (M = 18.66 years, SD = 1.01 years) and 2B had a sample of 72 (M = 18.56 years, SD = 0.82 years). The sample sizes were set to approximate the sample sizes used in Experiment 1A and 1B. 129 participants were recruited for Experiment 2A, with 11 excluded for the response training task and 6 excluded for the main task. An additional 34 participants were excluded for not noticing the change in the retro-cue (see below for details). 170 participants were recruited for Experiment 2B, with 42 excluded for the response training task, 12 for the main task, and 44 for the retro-cue. None of the recruited participants participated in other experiments in this paper.

In the main task, participants were instructed that the relevant feature on each trial would be indicated by an always-valid word cue. This retro-cue (duration = 500ms) appeared immediately after the memory display and always validly indicated the to-be-probed feature. The structure of the surprise task paradigm was otherwise retained such that the critical feature was first cued and probed only in the 26th trial (see Figure 1B). To filter out participants who did not pay attention to the retro-cue, we included an end-of-experiment survey in which participants responded whether they noticed that the word cue switched from “COLOR” to “LOCATION” (Experiment 2A) or “ORIENTATION” (Experiment 2B). Participants who reported not noticing the word cue change were then excluded from further analyses.

As Experiment 2A and 2B were not replication experiments, the Bayes factors, where applicable, were computed with the default prior settings in JASP. Additional analyses were carried out to compare the effects from these experiments to those in Experiment 1A and 1B respectively.

Results

Effects of Task Relevance

Despite the cue indicating the to-be-reported feature being presented immediately after the stimulus, data from experiment 2A showed once again that location memory improved significantly from SUR to FPS (Figure 2C). The mean absolute error in SUR was 45.97° (SE = 5.36°), which decreased to 9.30° in FPS (SE = 1.54°), BF_{10} = 7.33E+06. To compare these results with experiment 1A, we performed a mixed Bayesian ANOVA with experiment (between-subject: 1A or 2A) and trial type (within-subject: SUR or FPS) and found the BF_{incl} for the interaction term to be 0.73, indicating no strong evidence on whether the memory improvement was different between Experiment 1A and 2A. Consistent with the decrease in mean absolute error, the precision parameter s decreased from 22.01° (95% CI [13.46, 39.81]) to 8.39° (95% CI [7.03, 10.47]). The probability of guessing, P_U, also decreased from 0.44 (95% CI [0.24, 0.60]) to 0.06 (95% CI [0.02, 0.14]). The Kolmogorov-Smirnov test shows that location memory in SUR was significantly different from a uniform distribution, D = 0.24, $p < .001$. The evidence suggests that the amount of categorical bias was similar between trials, as the distance from the nearest canonical value in SUR (26.76°, SE = 2.34°) was numerically but not significantly smaller than that in FPS (31.70°, SE = 2.11°), BF_{10} = 0.332.

Experiment 2B shows that orientation memory also improved significantly from SUR to FPS (Figure 2D). The mean absolute error decreased from 79.61° (SE = 5.84°) to 25.18° (SE = 3.69°), BF_{10} = 2.01E+09. Similarly, we compared this memory improvement with that in

Experiment 1B using a mixed Bayesian ANOVA which suggested that the amount of improvement was indeed similar across experiments ($BF_{incl} = 0.17$ for the interaction term). Parameter s decreased from 114.57° (95% CI [10.51, 95.94]) to 15.88° (95% CI [12.42, 23.19]). The probability of guessing P_U, however, increased from 1.18E-13 (95% CI [0.34, 1.00]) to 0.21 (95% CI [0.10, 0.35]). For both parameters, the CI for SUR was wide such that it overlaps with the MLE for FPS, although the Kolmogorov-Smirnov test shows that SUR was again indistinguishable from pure guessing, $D = 0.12$, $p = .197$. Evidence again suggests that categorical bias did not differ between trials, distance from the nearest canonical value in SUR: 23.38° (SE = 1.7°), in FPS: 26.01° (SE = 1.81°), $BF_{10} = 0.183$.

Comparing Experiment 2A and 2B, the improvement in memory from SUR to FPS was numerically greater for orientation than location, although a mixed Bayesian ANOVA did not provide strong evidence for or against including the interaction effect between critical feature and trial type to be included in the model, $BF_{incl} = 1.39$. See Figure 3 (dotted lines).

Effect of Resource Recruitment

In Experiment 2A, evidence shows that color memory *did not* grow worse from Pre to Post (Figure 4C). The mean absolute error in Pre was 13.96° (SE = 0.53°) and that in Post was 14.65° (SE = 0.68°), $BF_{10} = 0.15$. The estimated precision in Pre was 15.24° (SE = 0.44°) and that in Post was 15.00° (SE = 0.71°), $BF_{10} = 0.21$. P_U in Pre was 0.038 (SE = 0.009) and that in Post was 0.044 (SE = 0.007), $BF_{10} = 0.31$. However, the numerical values were very similar to those in Experiment 1A, and mixed Bayesian ANOVAs show that the BF_{incl} for the interaction effects between experiment (between-subject: 1A or 2A) and trial type (within-subject: Pre or Post) are in general in favor of the Pre-Post cost being similar between experiments (mean absolute error: $BF_{incl} = 0.54$, s: $BF_{incl} = 0.32$, P_U: $BF_{incl} = 0.20$). We reason that this discrepancy might be because of the fact that replication BFs were used for Experiments 1 but default prior BFs were used for Experiments 2. As the numerical resource recruitment cost from location encoding is quite small, the data might only be detectable as positive evidence in the case of Experiment 1A where the prior distribution was adjusted to be the posterior distribution from an initial experiment that found a similarly small effect size. We do not think it appropriate to apply replication BFs to analyze the data from Experiments 2 because their designs are different from the previous experiments, such that further research is required to figure out whether location encoding induces a cost or not.

On the other hand, generally positive evidence was found for a memory cost to color in the Pre-Post comparison in Experiment 2B (Figure 4D). The mean absolute error for color increased from 13.63° (SE = 0.55°) to 16.89° (SE = 0.89°), $BF_{10} = 28.73$. The evidence is rather ambiguous on whether parameter s and P_U changed: s in Pre was 15.32° (SE = 0.51°) and in Post it was 16.32° (SE = 0.65°), $BF_{10} = 0.47$; P_U in Pre was 0.033 (SE = 0.010) and in Post it was 0.052 (SE = 0.007), $BF_{10} = 0.85$. Comparing these effects with those in Experiment 1B with a mixed Bayesian ANOVA, we again found evidence generally in favor of the Pre-Post cost being similar across experiments based on the BF_{incl} of the interaction term (mean absolute error: $BF_{incl} = 0.18$, s: $BF_{incl} = 0.38$, P_U: $BF_{incl} = 0.19$).

To compare Experiment 2A and 2B, a mixed Bayesian ANOVA with critical feature (between-subject: location or orientation) and trial type (within-subject: Pre or Post) was performed for each measurement (mean absolute error, s and P_U). Using s and P_U, the BF_{incl} of the interaction term suggests that the memory cost to color was comparable in Experiment 2A and 2B (s: $BF_{incl} = 0.23$, P_U: $BF_{incl} = 0.22$). The evidence was less clear when mean absolute error was used, $BF_{incl} = 1.31$.

Discussion

With the retro-cue, the critical feature was cued for report immediately after the stimulus disappeared. Therefore, we were able to obtain memory reports of the critical feature on the surprise trial that reflect how the feature was initially encoded with minimal influence from post-encoding processes such as inhibition and maintenance that can selectively damage the memory of features that are thought to be less relevant. That Experiment 2A and 2B still found significant task relevance effects suggest that the effect of task relevance was at least partly contributed by differences in encoding: Evidence suggests that the relevance effect was comparable between the two orientation experiments (1B and 2B), indicating that orientation encoding was the primary factor in causing the task relevance effect in Experiment 1B as well. Still, the between-experiment comparison for the location experiments (1A and 2A) did not find conclusive evidence in either direction. Further studies would be required to investigate whether post-encoding processes indeed also influence location memory when it is task relevant versus irrelevant.

The retro-cue also allows us to specifically measure the cost of additionally encoding the critical feature, as the maintenance load was the same between the Pre and Post color trials. Experiment 2A suggests that encoding location was not costly, while Experiment 2B suggests that encoding orientation was costly. However, evidence also suggests that the cost observed in each of these experiments was not different from that observed from the corresponding Experiment 1. In other words, the evidence is generally in favor of orientation encoding requiring additional resources, but further studies would be required to investigate whether location encoding is cost-free or not.

A remaining possibility is that participants encoded a highly accurate representation of the critical feature throughout the experiment, but on the surprise trial, its memory was disrupted by the inherent surprising elements of the surprise trial, namely, the first presentation of a new word cue and a new response screen (although the response screen was previously used in the training task). This could explain the task relevance effects in the previous experiments if it was assumed that these surprising elements led to a significant disruption to the memory representations in the surprise trial.

Experiment 3 was carried out to control for the potential memory loss induced by this surprising event. In Experiment 3, location was always the task relevant feature, but the response method was indicated by a retro-cue. Participants were cued to use one response method for the first 25 trials and then were surprised with a new word cue and a new response method on the 26th trial. If the surprising appearance of the cue and the response method per se can account for the task relevance effect, the SUR-FPS comparison should yield a comparable memory cost to Experiment 2A. On the other hand, if the surprise trial in Experiment 3 shows little-to-no memory reduction, it would support our postulation that the task relevance costs in the previous experiments reflect a lack of full automaticity in the feature encoding process.

Experiment 3

Methods

Unless otherwise specified, Experiment 3 used the same methods as Experiment 2A. The sample size for Experiment 3 was 82 (M = 18.76 years, SD = 0.94 years). The sample size was again set to approximate the one used in the previous experiments. 121 participants were recruited, with 3 excluded for the response training task, 11 excluded for the main task, and 25 excluded for the retro-cue.

In this experiment, location was always the probed feature, but participants were sometimes required to report it with the "wait" method, sometimes with the "drag" method (the drag method is identical to the one used in Experiment 1A and 1B for location report). The response screen of the wait method has the same basic components as that of the drag method, i.e., the black X and the outline circle. Instead of having its position being controlled by the cursor position ("dragged"), the black X will move spontaneously along the outline circle, starting from a random location at a speed of 0.3° per frame anticlockwise. The X would stop at its current position when a mouse click was detected. As in the drag method, they were given unlimited time and chances to respond.

Participants first completed two blocks of 10 perceptual matching trials, using one of the two response methods in each block (block order was randomized). Then, in the main task (Figure 1C), participants responded with the wait method for the first 25 before responding with the drag method in the 26th trial. This was followed by 24 control trials where the wait and drag method was cued and used with equal chance. The retro-cue would indicate the response method on that trial. The end-of-experiment survey queried whether participants noticed that the retro-cue switched from "WAIT" to "DRAG", and participants who reported not noticing this switch were excluded from further analyses.

SUR-FPS comparison was done on the surprise and the first post-surprise drag trial to assess the amount of memory loss due to an unexpected change in the retro word cue and the response method. The Pre-Post comparison was omitted. Additional analyses were carried out to compare the effects from this Experiment to that from Experiment 2A.

Results

Experiment 3 shows only limited evidence that location reports with the drag method improved from SUR to FPS (Figure 2E). The mean absolute error numerically decreased from 8.01° in SUR (SE = 1.7°) to 5.47° in FPS (SE = 0.52°), but it is ambiguous whether there was an evident decrease in the error or not ($BF_{10} = 0.65$). The model estimate s suggests that the precision essentially stayed the same, with s being 6.87° in SUR (95% CI [5.60, 9.11]) and 7.19° in FPS (95% CI [6.13, 8.31]). P_U decreased from 0.047 in SUR (95% CI [0.013, 0.133]) to 3.62E-14 in FPS (95% CI [2.98E-14, 0.0604]), although the SUR estimate is still within the CI for the FPS estimate and the CI is reasonably narrow. The Kolmogorov-Smirnov test also shows that SUR was distinct from a uniform distribution, D = 0.43, $p < .001$. Analyses further found evidence against a change in categorical bias between trials, distance from the nearest canonical value in SUR: 32.25° (SE = 2.34°), in FPS: 30.87° (SE = 1.43°), $BF_{10} = 0.141$.

The extent of the surprise memory loss was compared with that in Experiment 2A with a mixed Bayesian ANOVA. Importantly, the interaction term between experiment (between-subject: 2A or 3) and trial type (within-subject: SUR or FPS) was strongly suggested to be included in the model (BF_{incl} = 1.22E+07), see also Figure 3 (dashed line). The surprise memory loss in Experiment 3, if any, was significantly smaller than that in Experiment 2A.

Discussion

In Experiment 3 where subjects expected to remember location in the pre-surprise trials, there was barely any difference in the SUR-FPS comparison even though they were given an abrupt change in the word cue and response method on the surprise trial. The data therefore suggests that the task relevance effects observed in Experiments 1 and 2 should not be primarily attributed to the disruptions from surprising elements inherent to the surprise trial. A more likely explanation for the task relevance effect is that the encoding process was not as effective when the feature was considered task irrelevant. In other words, this suggests that the encoding of location and orientation were both not fully automatic.

General Discussion

We assessed the automaticity of location and orientation memory in VWM across five experiments. Experiment 1A and 1B show that location and orientation both exhibited reduced accuracy when probed on a surprise trial, and the additional requirement to remember each of these features recruited a significant amount of cognitive resource as measured in the comparison between color reports in pre-surprise and post-surprise trials. Consistent with previous literature, location demonstrated a higher degree of automaticity than orientation by showing a smaller surprise memory cost and recruiting a smaller amount of resource. However, location encoding was evidently not completely automatic. To study whether the incomplete automaticity in these features can be attributed to incomplete automaticity in their encoding (instead of post-encoding processes such as active inhibition and selective maintenance), Experiment 2A and 2B introduced always-valid feature retro-cues the made the critical feature immediately task relevant after the offset of the memory displays. Again, both location and orientation suffered a surprise memory cost, but that of location was numerically although not evidently smaller. On the other hand, while orientation was again found to be cognitively costly, the resource recruitment effect was not found for location. Experiment 3 tested the alternative explanation that the poor surprise trial memory was due to disruptions from the surprising elements in that trial. In contrast to

Experiments 1 and 2, now there is no evidence that memory was worse on the surprise trial than the first post-surprise trial, suggesting that the task irrelevance effects observed were unlikely to be due to surprise-related disruptions.

In summary, although location encoding demonstrated greater automaticity to orientation encoding, neither of these encoding processes was fully automatic. We note here that the incomplete automaticity can be attributed to incomplete automaticity in the perceptual encoding process, which is thought to be a fast process that builds a rather short-lived perceptual representation (Potter, 1976, 1993), or the working memory encoding process, which is thought to be relatively slower but can result in a durable, readily accessible representation (Ricker & Hardman, 2017) or both. Further studies would be required to differentiate these possibilities.

Automaticity in Location Encoding

Evidence from Experiment 1A, 2A, and 3 together support the notion that location is not fully automatically encoded into VWM. It was also consistently shown that some, but not all location information was incidentally encoded, as location report on the surprise trial was still distinctly clustered around the true value despite being worse than that on the first post-surprise trial. This pattern suggests *partial automaticity* in location encoding, which is in line with a graded view of automaticity (Hasher & Zacks, 1979; Logan, 1985). Still, it seems inconsistent with previous studies that explicitly probed location memory when it was task irrelevant and observed no cost (Chen & Wyble, 2015a, 2016, 2018; Mandler et al., 1977; McCormack, 1982; Pezdek & Evans, 1979; Schulman, 1973; Smith & Milner, 1989; Zechmeister et al., 1975). On this issue, we note that the current findings were based on a surprise trial that occurred after 25 color trials. While automaticity can be developed through practice (automatization, e.g., Shiffrin & Schneider, 1977), some form of *de-automatization* may also occur over practice, such that a test of a task irrelevant feature yields worse memory when it is presented later in the experiment (Chen & Wyble, 2018; Wyble et al., 2019). Therefore, the currently observed partial automaticity in location encoding may be pre-conditioned on some degree of *de-automatization*, which potentially explains the inconsistency with studies that probed task irrelevant location memory on the first trial (Mandler et al., 1977; McCormack, 1982; Pezdek & Evans, 1979; Schulman, 1973; Smith & Milner, 1989; Zechmeister et al., 1975).

In addition, in attempt to reconcile the current findings with the previous studies that used a surprise trial paradigm to probe location, we conducted a meta-analysis on the relevant experiments in these previous studies (Chen & Wyble, 2016, 2018; see supplemental materials for details). When each experiment was examined individually, the SUR performance was numerically but not significantly worse than that of FPS. However, when data from these studies were considered altogether, a paired Bayesian t-test between SUR and FPS accuracy in fact revealed substantial evidence against the null ($BF_{10} = 5.08$). Summing the number of participants who produced a correct location response in each condition across experiments, we found the surprise trial performance to be significantly worse than that of the next control trial, $\chi^2(1, N = 110) = 5.88$, $p = .0153$, but critically, also significantly higher than the 25% chance level, $\chi^2(1, N = 110) = 165.93$, $p < .001$. This meta-analysis shows that in previous surprise trial studies, incidentally encoded location was neither comparable to intentionally encoded location nor to the chance level, compatible with the findings from Experiment 1A and 2A in this paper.

A recent study by Woodman (2021) corroborates with our data in showing that location encoding is not fully automatic. His study shows that having to remember both location and color worsened change detection performance compared to when only either of the features had to be remembered. Furthermore, this memory cost was observed when the stimulus-onset asynchrony between the memory items and the mask was as short as 35ms, strongly suggesting that the cost was from the added requirement to intentionally *encode* location. It is possible that we did not find a cost to color memory when location was additionally encoded (Experiment 2A) due to the small effect size (we elaborated on this point in the Results section of Experiment 2A and 2B).

While the current study focused on showing the incompleteness of location automaticity, we still found higher automaticity for location than orientation in the surprise trial. This is generally consistent with the idea of a location "privilege" (e.g., Pertzov & Husain, 2014). However, the relatively precise location reports in the surprise trial may have been supported by spatial attention, as previous studies have found that the task irrelevant location of a memory target can be decoded from alpha oscillations, which have been associated with covert spatial attention (Bae & Luck, 2018; Foster et al., 2017). Thus, further investigations will be needed to see whether the location privilege in a surprise memory trial will remain if spatial attention is distracted from the memory target.

Automaticity in Orientation Encoding

We failed to observe any trace of automaticity in orientation encoding. Particularly, the surprise orientation reports from Experiment 1B and 2B were indistinguishable from a uniform distribution. This finding was inconsistent with the orientation surprise trial data in Swan et al. (2016), which was clearly clustered around the true value (see Figure 5). We think that the critical distinction between our experiments and theirs is whether the orientation could easily be inferred by the location. Swan et al. (2016) used an arrow that always appeared at the center of the screen, such that the arrowhead had a distinct location that covaried with orientation. Relatedly, uninformative arrow cues have been shown to exogenously produce attentional shift towards the pointed-at location (Stevens et al., 2008). Thus, participants in Swan et al. (2016) may have had an incidental memory of the location of the arrowhead and/ or the pointed-at location and used that information to respond to the surprise orientation question. Our finding that location enjoys some automaticity supports this possibility. In the current study, each stimulus appeared at a random location, and orientation information was communicated through local variation of the stimulus. The separation of orientation and location information is supported by findings that orientation can be decoded from neural sources distinct from those that allow decoding of spatial location (Bae & Luck, 2018). Thus, our design minimizes the possibility of inferring orientation from global location memory. When orientation and location were thus disentangled in these experiments, there was no evidence of incidentally encoded orientation information.

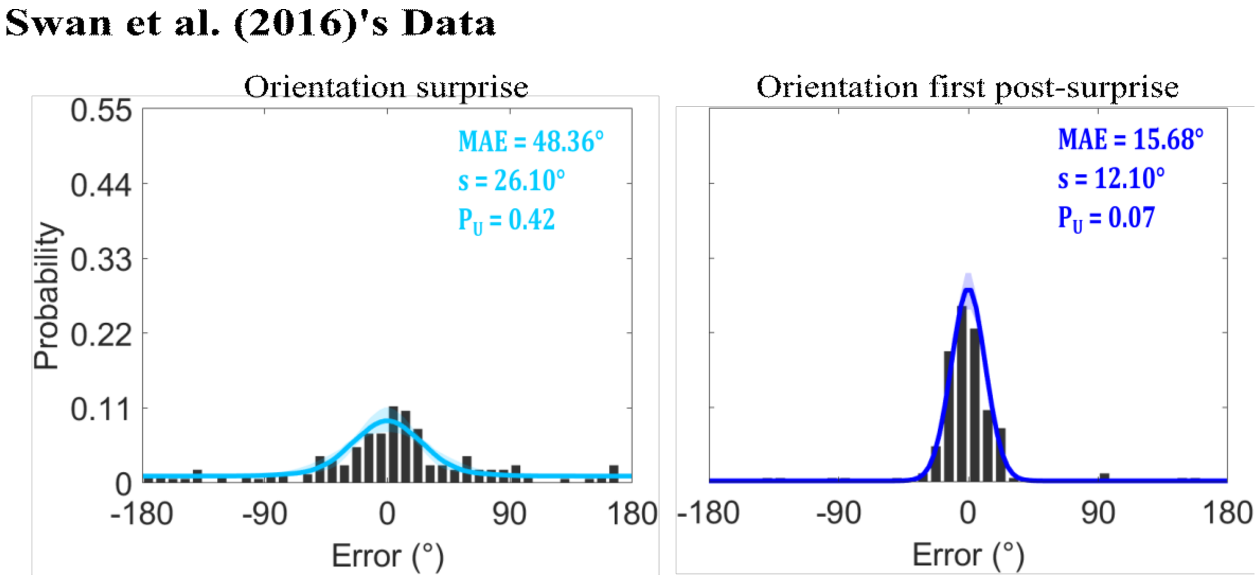


Figure 5. Orientation surprise and first post-surprise performance in Swan et al. (2016). Consistent with the current findings, a dramatic improvement was observed from the surprise to the first post-surprise trial. However, a measurable amount of orientation information was reportable by participants in Swan et al. (2016) on the surprise trial, while surprise orientation report in the current study (Experiment 1B and 2B) was indistinguishable from a uniform distribution. Shaded areas are 95% credible intervals. The figures are adapted from Figure 6 in the original article.

While the current orientation surprise data is consistent with previous observations of near-chance performance when other task irrelevant nonspatial attributes were explicitly probed (Chen & Wyble, 2015a, 2015b, 2016, 2018), it appears at odds with previous studies showing that incidentally encoded nonspatial attributes still influenced the ongoing tasks (Bocincova et al., 2017; Ecker et al., 2013; Foerster & Schneider, 2018, 2020; Gao et al., 2016; Gao et al., 2010; Hollingworth & Luck, 2009; Hollingworth et al., 2013a, 2013b; Hyun et al., 2009; Logie et al., 2011; Soto & Humphreys, 2009; Yin et al., 2012; Yin et al., 2011). Indeed, these studies present strong evidence that some nonspatial attributes are at least partially incidentally encoded into VWM, and we also do not wish to claim that the encoding of all nonspatial attributes are completely void of automaticity. Instead, it is possible that the encoding of orientation information communicated through local stimulus variation (as in the current study) is particularly effortful (Becker et al., 2013; Gao et al., 2010; Miller et al., 2014). In particular, using Landolt-Cs with varying colors and orientations in a change detection task, Gao et al. (2010) found that change detection accuracy was impeded by a task irrelevant color change but not by a task irrelevant orientation change. Furthermore, it was shown that immediate backward masking (like the design used in the current study) might be instrumental to prevent the encoding of an irrelevant feature (Gu et al., in press). Therefore, further studies are required to investigate whether local orientation information is incidentally encoded under other circumstances.

Partial Automaticity

We posited that an intermediate pattern in the SUR-FPS comparison where SUR is neither comparable to FPS nor to the chance level would be more consistent with a graded view of automaticity, but we should also consider the possibility that automaticity is probabilistically dichotomous. Namely, incidental encoding may result in intentional-comparable memory in a certain proportion of trials and no memory in the rest. While our experiments were not designed to directly address this possibility, we did observe that from SUR to FPS, not only did the guess rate decreased, but the precision also improved. The improvement in precision is more in line with the idea that automaticity can indeed be graded and preliminarily suggests that the memory data we observed from the surprise trials reflect *partial automaticity* in encoding. However, more research would be required to understand the nature of incomplete automaticity under different circumstances.

Several possible manifestations of partial automaticity might underlie our results. For example, Zbrodoff & Logan (1986) suggested that task relevance can have distinct effects on a process' initiation, continuation, and termination. A possibility is that location encoding is automatically initiated but requires intention to be run to completion, which explains the current finding where some location information was incidentally encoded but making location task relevant further improved its precision. In addition, a process need not be completely dependent or independent from cognitive resources and that automaticity can vary with the relative level of independence between the two (Moors & De Houwer, 2006). The current data, however, does not make a strong case on whether location encoding indeed recruits a significant amount of resource.

Another line of research arguing against binary automaticity modes has shown that primary features of automaticity do not always co-occur (Bargh, 1992; Kahneman & Chajczyk, 1983). For example, naming the ink color of a word is slowed for an incongruent color word (the Stroop effect), suggesting that reading is a mandatory process (e.g., Keele, 1972). However, Kahneman & Chajczyk (1983) showed that the Stroop Effect can be moderated by re-allocation of attention, such that the cost of incongruency was reduced when a black, non-color word was presented in proximity to the color word. In other words, reading occurs without intention but can be interfered with resource re-allocation, thus fitting one primary feature of automaticity but not another. This is also exemplified in our experiments. In Experiment 1A and 1B, both location and orientation were affected by task relevance, and both incurred a small cost on color, such that the two characteristics of automaticity are generally in agreement. While this pattern was retained for orientation encoding in Experiment 2B, Experiment 2A suggests that location encoding is influenced by task relevance but does not induce a clear cost on other ongoing cognitive processes.

Encoding Automaticity Constrains Organization of VWM Representations

The automaticity of feature encoding has implications on how representations can possibly be organized. In visual working memory, incidental encoding of a feature seems to be a pre-requisite for a feature to be the “index” of the memory items, i.e., the feature based on which representations are organized, and from which other information about the item can be accessed. VWM has often been proposed to rely at least partly on a spatial reference frame (Bays et al., 2009; Bays, 2014; Swan & Wyble, 2014; Oberauer & Lin, 2017), which implies that some degree of automaticity in location encoding is required. Our findings are generally consistent with this assumption but also show that location encoding is not a fully automatic process in that incidentally encoded location information is relatively imprecise. This implies that the location resolution of spatial reference frames in VWM models, if present, would be somewhat coarse by default. More specific indexing of representations may require aids from other information (e.g., temporal indexing, see Harrison & Bays, 2018; Heuer & Rolfs, 2021; Schneegans & Bays, 2019; Swan & Wyble, 2014). However, a caveat is that only one item was presented per trial in the current study, posing little to no demand on feature binding. It is possible that location encoding is only necessary for binding, and participants in our experiment 1A and 2A were able to make use of “free-floating” color representations to support their memory reports before the location surprise trial (see Treisman & Gelade, 1980; Treisman et al., 1992). Although memory of features such as orientation enjoys high fidelity at set size one (Magnussen, 2000; Wilken & Ma, 2004), a counterintuitive prediction from this account is that the *relative fidelity of location* memory on the surprise trial compared to the first post-surprise trial would actually improve as set size increases, and further studies would be required to investigate this possibility.

The finding in experiment 1B and 2B that orientation encoding had no trace of automaticity is consistent with the emerging literature suggesting that VWM is not constrained by a strict item-based limit and that representations are not necessarily object-based (Fougnie & Alvarez, 2011). In the context of feature-based theories, the absence of orientation information on the surprise trial is consistent with the findings that different nonspatial features can be remembered independently (Bays et al., 2011; Fougnie & Alvarez, 2011). Furthermore, the cost induced to color memory after orientation became task relevant corroborates previous studies showing that memory performance depends on item complexity (Alvarez & Cavanagh, 2004) and the number of task relevant features (Cowan et al., 2013; Fougnie et al., 2010; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013; Swan et al., 2016), but is inconsistent with studies showing that memory of different features draws on independent resources (Magnussen, 2000; Wang et al., 2017).

Summary

Across five experiments, we showed that location encoding is more automatic than orientation encoding, but location encoding is still not fully automatic. Our findings constrain the resolution of location indexing in space-based representational theories of VWM and are against object-based encoding theories. The demonstrated partial automaticity in location encoding is consistent with conceptualizations of automaticity as a continuum instead of binary modes.

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*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Arnold, T. B., & Emerson, J. W. (2011). Nonparametric goodness-of-fit tests for discrete null distributions. *R Journal*, *3*(2), 34–39. https://doi.org/10.32614/rj-2011-016

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126. https://doi.org/10.1016/S1364-6613(00)01593-X

Awh, E., Jonides, J., Smith, E. E., Buxton, R. B., Frank, L. R., Love, T., … Gmeindl, L. (1999). Rehearsal in Spatial Working Memory: Evidence from Neuroimaging. *Psychological Science*, *10*(5), 433–437. https://doi.org/10.1111/1467-9280.00182

Bae, G. Y., & Luck, S. J. (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *Journal of Neuroscience*, *38*(2), 409–422. https://doi.org/10.1523/JNEUROSCI.2860-17.2017

Bae, G. Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General*, *144*(4), 744–763. https://doi.org/10.1037/xge0000076

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105*, 181–199.

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *Journal of Neuroscience*, *34*(10), 3632–3645. https://doi.org/10.1523/JNEUROSCI.3204-13.2014

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. https://doi.org/10.1167/9.10.7.Introduction

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*(6), 1622–1631. https://doi.org/10.1016/j.neuropsychologia.2010.12.023

Becker, M. W., Miller, J. R., & Liu, T. (2013). A severe capacity limit in the consolidation of orientation information into visual short-term memory. *Attention, Perception, and Psychophysics*, *75*(3), 415–425. https://doi.org/10.3758/s13414-012-0410-0

Bocincova, A., van Lamsweerde, A. E., & Johnson, J. S. (2017). The role of top-down suppression in mitigating the disruptive effects of task-irrelevant feature changes in visual working memory. *Memory and Cognition*, *45*(8), 1411–1422. https://doi.org/10.3758/s13421-017-0738-8

Chen, H., & Wyble, B. (2015a). The location but not the attributes of visual cues are automatically encoded into working memory. *Vision Research*, *107*, 76–85. https://doi.org/10.1016/j.visres.2014.11.010

Chen, H., & Wyble, B. (2015b). Amnesia for Object Attributes: Failure to Report Attended Information That Had Just Reached Conscious Awareness. *Psychological Science*, *26*(2), 203–210. https://doi.org/10.1177/0956797614560648

Chen, H., & Wyble, B. (2016). Attribute Amnesia Reflects a Lack of Memory Consolidation for Attended Information. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(2), 225–234. https://doi.org/10.1037/xhp0000133

Chen, H., & Wyble, B. (2018). The neglected contribution of memory encoding in spatial cueing: A new theory of costs and benefits. *Psychological Review*, *125*(6), 936–968. https://doi.org/10.1037/rev0000116

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170–178. https://doi.org/10.1016/S1364-6613(00)01476-5

Chun, M. M., & Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, *36*, 28–71. https://doi.org/10.1109/CICN.2016.58

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N., & Morey, C. C. (2007). How Can Dual-Task Working Memory Retention Limits Be Investigated? *Psychological Science*, *18*(8), 686–688. https://doi.org/10.1111/j.1467-9280.2007.01960.x

Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to Attributes and Objects in Working Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(3). https://doi.org/10.1038/jid.2014.371

De Vries, I. E. J., Van Driel, J., Karacaoglu, M., & Olivers, C. N. L. (2018). Priority Switches in Visual Working Memory are Supported by Frontal Delta and Posterior Alpha Interactions. *Cerebral Cortex*, *28*(11), 4090–4104. https://doi.org/10.1093/cercor/bhy223

Ecker, U. K. H., Maybery, M., & Zimmer, H. D. (2013). Binding of intrinsic and extrinsic features in working memory. *Journal of Experimental Psychology: General, 142*, 218–234. http://dx.doi.org/10.1037/ a0028732

Eitam, B., Yeshurun, Y., & Hassan, K. (2013). Blinded by irrelevance: Pure irrelevance induced "blindness." *Journal of Experimental Psychology: Human Perception and Performance*, *39*(3), 611–615. https://doi.org/10.1037/a0032269

Ester, X. E. F., Sprague, X. T. C., & Serences, X. J. T. (2020). Categorical Biases in Human Occipitoparietal Cortex. *The Journal of Neuroscience*, *40*(4), 917–931.

Foerster, R. M., & Schneider, W. X. (2018). Involuntary top-down control by search-irrelevant features : Visual working memory biases attention in an object-based manner. *Cognition*, *172*(May 2017), 37–45. https://doi.org/10.1016/j.cognition.2017.12.002

Foster, J. J., Bsales, E. M., Jaffe, R. J., & Awh, E. (2017). Alpha-Band Activity Reveals Spontaneous Representations of Spatial Position in Visual Working Memory. *Current Biology*, *27*(20), 3216-3223.e6. https://doi.org/10.1016/j.cub.2017.09.031

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), 1–12. https://doi.org/10.1167/11.12.3.Introduction

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision*, *10*(27), 1–11. https://doi.org/10.1167/10.12.27.Introduction

Gao, Z., Li, J., Yin, J., & Shen, M. (2010). Dissociated mechanisms of extracting perceptual information into visual working memory. *PLoS ONE*, *5*(12), 1–15. https://doi.org/10.1371/journal.pone.0014273

Gao, Z., Yu, S., Zhu, C., Shui, R., Weng, X., Li, P., & Shen, M. (2016). Object-based Encoding in Visual Working Memory: Evidence from Memory-driven Attentional Capture. *Scientific Reports*, *6*(March), 1–7. https://doi.org/10.1038/srep22822

Gu, Q., Dai, A., Ye, T., Huang, B., Lu, X., Shen, M., & Gao, Z. (in press). Object-based encoding in visual working memory: A critical revisit. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218211052502

Gu, Q., Wan, X., Ma, H., Lu, X., Guo, Y., Shen, M., & Gao, Z. (2020). Event-based encoding of biological motion and location in visual working memory. *Quarterly Journal of Experimental Psychology*, *73*(8), 1261–1277. https://doi.org/10.1177/1747021820903042

Hakim, N., Adam, K. C. S., Gunseli, E., Awh, E., & Vogel, E. K. (2019). Dissecting the Neural Focus of Attention Reveals Distinct Processes for Spatial Attention and Object-Based

Storage in Visual Working Memory. *Psychological Science*, *30(4)*, 526–540. https://doi.org/10.1177/0956797619830384

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*(2), 325–347. https://doi.org/10.1037/xlm0000031

Harms, C. (2019). A Bayes Factor for Replications of ANOVA Results. *American Statistician*, *73*(4), 327–339. https://doi.org/10.1080/00031305.2018.1518787

Harrison, W. J., & Bays, P. M. (2018). Visual working memory is independent of the cortical spacing between memoranda. *Journal of Neuroscience, 38*(12), 3116–3123. https://doi.org/10.1523/JNEUROSCI.2645-17.2017

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356–388.

Heuer, A., & Rolfs, M. (2021). Incidental encoding of visual information in temporal reference frames in working memory. *Cognition, 207*(May 2020), 104526. https://doi.org/10.1016/j.cognition.2020.104526

Hollingworth, A., & Bahle, B. (2020). Feature-based guidance of attention by visual working memory is applied independently of remembered object location. *Attention, Perception, & Psychophysics*, *82*, 98–108.

Hollingworth, A., & Luck, S. J. (2009). The role of visual working memory (VWM) in the control of gaze during visual search. *Attention, Perception, & Psychophysics*, *71*(4), 936–949. https://doi.org/10.3758/APP

Hollingworth, A., Matsukura, M., & Luck, S. J. (2013a). Visual Working Memory Modulates Rapid Eye Movements to Simple Onset Targets. *Psychological Science*, *24*(5), 790–796. https://doi.org/10.1177/0956797612459767

Hollingworth, A., Matsukura, M., & Luck, S. J. (2013b). Visual working memory modulates low-level saccade target selection : Evidence from rapidly generated saccades in the global effect paradigm. *Journal of Vision*, *13(13)*(4), 1–18. https://doi.org/10.1167/13.13.4.doi

Hyun, J. Seok, Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The Comparison of Visual Working Memory Representations With Perceptual Inputs. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1140–1160. https://doi.org/10.1037/a0015019

JASP Team (2020). JASP (Version 0.14) [Computer software].

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of Visual Short-Term Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*(3), 683–702. https://doi.org/10.1037/0278-7393.26.3.683

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(4), 497–509. https://doi.org/10.1037//0096-1523.9.4.497

Keele, S. W. (1972). Attention demands of memory retrieval. *Journal of Experimental Psychology, 93*, 245- 248.

Levillain, F., & Flombaum, J. I. (2012). Correspondence problems cause repositioning costs in visual working memory. *Visual Cognition*, *20*(6), 669–695. https://doi.org/10.1080/13506285.2012.683050

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology, 39*, 367–386.

Logie, R. H., Brockmole, J. R., & Jaswal, S. (2011). Feature binding in visual short-term memory is unaffected by task-irrelevant changes of location, shape, and color. *Memory and Cognition*, *39*(1), 24–36. https://doi.org/10.3758/s13421-010-0001-z

Luck, S. J., & Vogel, E. K. (1997). The Capacity of Visual Working Memory for Features and Conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.3950/jibiinkoka.58.216

Ly, A., Etz, A., Marsman, M., & Wagenmakers, E. J. (2019). Replication Bayes factors from evidence updating. *Behavior Research Methods*, *51*(6), 2498–2508. https://doi.org/10.3758/s13428-018-1092-x

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences*, *23*(6), 247–251. https://doi.org/10.1016/S0166-2236(00)01569-1

Mandler, J. M., Seegmiller, D., & Day, J. (1977). On the coding of spatial information. *Memory & Cognition*, *5*(1), 10–16. https://doi.org/10.3758/BF03209185

McCormack, P. D. (1982). Coding of spatial information by young and elderly adults. *Journals of Gerontology*, *37*(1), 80–86. https://doi.org/10.1093/geronj/37.1.80

Miller, J. R., Becker, M. W., & Liu, T. (2014). The bandwidth of consolidation into visual short-term memory depends on the visual feature. *Visual Cognition*, *22*(7), 920–947. https://doi.org/10.1080/13506285.2014.936923

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

Oberauer, K. (2013). The focus of attention in working memory-from metaphors to mechanisms. *Frontiers in Human Neuroscience*, *7*(OCT), 1–16. https://doi.org/10.3389/fnhum.2013.00673

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory and Cognition*, *41*(8), 1212–1227. https://doi.org/10.3758/s13421-013-0333-6

Oberauer, K., & Lin, H. Y. (2017). An interference model of visual working memory. *Psychological Review*, *124*(1), 21–59. https://doi.org/10.1037/rev0000044

Oh, S. H., & Kim, M. S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin and Review*, *11*(2), 275–281. https://doi.org/10.3758/BF03196570

Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental*

Psychology: Human Perception and Performance, *32*(5), 1243–1265. https://doi.org/10.1037/0096-1523.32.5.1243

Palmer, J., Boston, B., & Moore, C. M. (2015). Limited capacity for memory tasks with multiple features within a single object. *Attention, Perception, and Psychophysics*, *77*(5), 1488–1499. https://doi.org/10.3758/s13414-015-0909-2

Park, Y. E., Sy, J. L., Hong, S. W., & Tong, F. (2017). Reprioritization of Features of Multidimensional Objects Stored in Visual Working Memory. *Psychological Science*, *28*(12), 1773–1785. https://doi.org/10.1177/0956797617719949

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.* https://doi.org/10.3758/s13428-018-01193-y

Pertzov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception, and Psychophysics*, *76*(7), 1914–1924. https://doi.org/10.3758/s13414-013-0541-y

Pilling, M., & Gellatly, A. (2013). Task probability and report of feature information: What you know about what you “see” depends on what you expect to need. *Acta Psychologica*, *143*(3), 261–268. https://doi.org/10.1016/j.actpsy.2013.04.007

Potter, M. C. (1978). Short-term Conceptual Memory for Pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(5), 509–522. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0096151507603168

Potter, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition*, *21*(2), 156–161. https://doi.org/10.3758/BF03202727

Pratte, M. S., Park, Y. E., Rademaker, R. L., & Tong, F. (2017). Accounting for Stimulus-Specific Variation in Precision Reveals a Discrete Capacity Limit in Visual Working Memory. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(1), 6–17.

Quak, M., Langford, Z. D., London, R. E., & Talsma, D. (2018). Contralateral delay activity does not reflect behavioral feature load in visual working memory. *Biological Psychology, 137*(July 2016), 107–115. https://doi.org/10.1016/j.biopsycho.2018.07.006

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rajsic, J., & Wilson, D. E. (2014). Asymmetrical access to color and location in visual working memory. *Attention, Perception, and Psychophysics*, *76*(7), 1902–1913. https://doi.org/10.3758/s13414-014-0723-2

Ricker, T. J., & Hardman, K. O. (2017). The nature of short-term consolidation in visual working memory. *Journal of Experimental Psychology: General*, *146*(11), 1551–1573. https://doi.org/10.1037/xge0000346

Ricker, T. J., Cowan, N., & Morey, C. C. (2010). Visual working memory is disrupted by covert verbal retrieval. *Psychonomic Bulletin & Review*, *17*(4), 516–521. https://doi.org/10.3758/PBR.17.4.516

Sala, J. B., & Courtney, S. M. (2009). Flexible working memory representation of the relationship between an object and its location as revealed by interactions with attention. *Attention, Perception, & Psychophysics*, *71*(7), 1525–1533. https://doi.org/10.3758/APP.71.7.1525.Flexible

Schneegans, S., & Bays, P. M. (2019). New perspectives on binding in visual working memory. *British Journal of Psychology, 110*(2), 207–244. https://doi.org/10.1111/bjop.12345

Schneegans, S., Harrison, W. J., & Bays, P. M. (2021). Location-independent feature binding in visual working memory for sequentially presented objects. *Attention, Perception, & Psychophysics*.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Schulman, A. I. (1973). Recognition memory and the recall of spatial location. *Memory & Cognition*, *1*(3), 256–260. https://doi.org/10.3758/BF03198106

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-020-00938-0

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Hernstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (Vol. 2, pp. 739–811). New York: Wiley.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Shin, H., & Ma, W. J. (2016). Crowdsourced single-trial probes of visual working memory for irrelevant features. *Journal of Vision*, *16(5)*(10), 1–8. https://doi.org/10.1167/16.5.10.doi

Shin, H., & Ma, W. J. (2017). Visual short-term memory for oriented, colored objects. *Journal of Vision*, *17(9)*(12), 1–19. https://doi.org/10.1167/17.9.12.doi

Smith, M. Lou, & Milner, B. (1989). Right hippocampal impairment in the recall of spatial location: Encoding deficit or rapid forgetting? *Neuropsychologia*, *27*(1), 71–81. https://doi.org/10.1016/0028-3932(89)90091-2

Soto, D., & Humphreys, G. W. (2009). Automatic Selection of Irrelevant Object Features Through Working Memory Evidence for Top-Down Attentional Capture. *Experimental Psychology*, *56*(3), 165–172. https://doi.org/10.1027/1618-3169.56.3.165

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory : 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, *78*, 1839–1860. https://doi.org/10.3758/s13414-016-1108-5

Stevens, S. A., West, G. L., Al-Aidroos, N., Weger, U. W., & Pratt, J. (2008). Testing whether gaze cues and arrow cues produce reflexive or volitional shifts of attention. *Psychonomic Bulletin & Review*, *15*(6), 1148–1153. https://doi.org/10.3758/PBR.15.6.1148

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision, 13*(10):9, 1–8. https://doi.org/10.1167/13.10.9

Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, and Psychophysics*, *76*(7), 2136–2157. https://doi.org/10.3758/s13414-014-0633-3

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), 1–12. https://doi.org/10.1167/16.3.32

Thayer, D. D., Bahle, B., & Hollingworth, A. (2021). Guidance of attention from visual working memory is feature-based, not object-based: Implications for models of feature binding. *Journal of Experimental Psychology: General,* http://dx.doi.org/10.1037/xge0001116

Treisman, A., & Gelade, G. (1980). A Feature-Integration of Attention. *Cognitive Psychology*, *12*, 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics*, *33*(6), 527–532. https://doi.org/10.3758/BF03202934

Van Lamsweerde, A. E., & Beck, M. R. (2011). The change probability effect: Incidental learning, adaptability, and shared visual working memory resources. *Consciousness and Cognition*, *20*(4), 1676–1689. https://doi.org/10.1016/j.concog.2011.09.003

Van Moorselaar, D., Theeuwes, J., & Olivers, C. N. L. (2014). In Competition for the Attentional Template: Can Multiple Items Within Visual Working Memory Guide Attention? *Journal of Experimental Psychology-Human Perception and Performance*, *40*(4), 1450–1464.

Verhagen, A. J., & Wagenmakers, E.-J. (2014). Bayesian tests to quantify the result of a replication attempt. *Journal of Experimental Psychology: General, 143*(4), 1457–1475.

Vidal, J. R., Gauchou, H. L., Tallon-Baudry, C., & O’Regan, J. K. (2005). Relational information in visual short-term memory: The structural gist. *Journal of Vision*, *5*(3), 244–256. https://doi.org/10.1167/5.3.8

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. https://doi.org/10.1038/nature02447

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*(1), 92–114. https://doi.org/10.1037/0096-1523.27.1.92

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. L., & Wang, Z. (2017). Separate Capacities for Storing Different Features in Visual Working Memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, *43*(2), 226–236. https://doi.org/10.1037/xlm0000295

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 1120–1135. https://doi.org/10.1167/4.12.11

Woodman, G. F. (2021). Spatial location is filtered out of visual working memory representations when task irrelevant, just like other features. *Attention, Perception, and Psychophysics*, *83*(4), 1391–1396. https://doi.org/10.3758/s13414-021-02263-8

Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, *11*(2), 269–274.

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2012). Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognition*, *20*(1), 1–28. https://doi.org/10.1080/13506285.2011.630694

Wyble, B., Hess, M., O'Donnell, R. E., Chen, H., & Eitam, B. (2019). Learning how to exploit sources of information. *Memory and Cognition*, *47*(4), 696–705. https://doi.org/10.3758/s13421-018-0881-x

Yin, J., Gao, Z., Jin, X., Ding, X., Liang, J., & Shen, M. (2012). The neural mechanisms of percept-memory comparison in visual working memory. *Biological Psychology*, *90*(1), 71–79. https://doi.org/10.1016/j.biopsycho.2012.02.023

Yin, J., Gao, Z., Jin, X., Ye, L., Shen, M., & Shui, R. (2011). Tracking the mismatch information in visual short term memory: An event-related potential study. *Neuroscience Letters*, *491*(1), 26–30. https://doi.org/10.1016/j.neulet.2011.01.001

Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General, 115*, 118–130.

Zechmeister, E. B., McKillip, J., Pasko, S., & Bespalec, D. (1975). Visual memory for place on the page. *Journal of General Psychology*, *92*, 43–52. https://doi.org/10.1080/00221309.1975.9711326

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860