Mixing and mingling: inter-item competition in visual working memory is both feature-general and feature-specific

Janna W. Wennberg[1], John T. Serences[1,2]

1. Department of Psychology, University of California, San Diego, La Jolla, CA
2. Neuroscience Graduate Program, University of California, San Diego, La Jolla, CA

Acknowledgements: We are grateful to Kirsten Adam for her input on experimental design and to Grayson Emery and Aisha Hill for their assistance with data collection. This study was supported by the National Eye Institute of the National Institute of Health via a grant awarded to John Serences (RO1-EY025872) and the University of California, San Diego. The authors have no conflicts of interest.

Correspondence should be addressed to Janna Wennberg (jwennber@ucsd.edu).

Abstract

Visual working memory (WM) is a central cognitive ability but is capacity-limited due to competition between remembered items. Understanding whether inter-item competition depends on the similarity of the features being remembered has important implications for determining if competition occurs in sensory or post-sensory stages of processing. Experiment 1 compared the precision of WM across homogeneous displays, where items belonged to the same feature type (e.g., colorful circles) and heterogeneous displays (e.g., colorful circles and oriented bars). Performance was better for heterogeneous displays, suggesting a feature-specific component of interference. However, Experiment 2 used a retro-cueing task to isolate encoding from online maintenance and revealed that inter-item competition during storage was not feature-specific. The data support recent models of WM in which inter-item interference—and hence capacity limits in WM—occurs in higher-order structures that receive convergent input from a diverse array of feature-specific representations.

Keywords: visual working memory, working memory interference, sensory recruitment hypothesis

Statement of Relevance

As we navigate the world, there is more information than we are able to process. This limitation is partially due to how little we are able to hold in working memory. Another feature of working memory is that we can hold many types of information in mind, from a phone number, to a face, to the color of a swatch of paint. Typically, holding multiple items in working memory leads to these items competing with each other for limited cognitive resources. Our study investigated the dynamics of this competition. Specifically, we tested whether competition is specific to the type of information being held (e.g., if colors only interfere with colors), or whether it is more general. Our results indicate that while competition in working memory is not feature-specific, feature-specific factors are, nevertheless, relevant.

Mixing and mingling: inter-item competition in visual working memory is both feature-general and feature-specific

Visual working memory, broadly defined as the ability to hold visual information "in mind," mediates many visually-guided behaviors, and is often disrupted in developmental and psychiatric disorders such as ADHD, Parkinson's, depression, and schizophrenia (Gold & Luck, 2023; Schecklmann et al., 2011). A critical feature of visual working memory is that it has limited capacity: most people cannot precisely remember details about more than three or four items (Adam et al., 2017; Alvarez & Cavanagh, 2004; Brady et al., 2011; Cowan, 2001; Liu & Jiang, 2005; Luck & Vogel, 1997; Ma et al., 2014). To date, these limitations can best be explained by *inter-item interference,* where multiple items in working memory compete for limited resources (Bays, 2014; Bays et al., 2009; Jonides & Nee, 2006; Lewis-Peacock & Norman, 2014; Ma et al., 2014; Oberauer & Lin, 2017). Indeed, there are often distortions of individual items in memory such that items are attracted towards or repelled from other items, highlighting the high degree of intermingling between representations (Chunharas et al., 2022; Golomb, 2015; Lively et al., 2021; Scotti et al., 2021; Wyble & Swan, 2015).

Many models of flexible information storage explicitly or implicitly suggest that inter-item interference arises due to competition between sensory representations, which is consistent with *sensory recruitment,* or a role for sensory neurons that encode specific features in supporting high-fidelity working memory for those features (Adam et al., 2022; Bays, 2014; D'Esposito & Postle, 2015; Ester et al., 2013; Gayet et al., 2018; Pratte & Tong, 2014; Schurgin et al., 2020; Serences et al., 2009; Sprague et al., 2014). Accordingly, some sensory recruitment models predict that feature similarity should be a determinant of inter-item competition: interference should be higher for more similar items that are encoded by overlapping neural populations. Behavioral studies generally suggest that competition is mediated by feature similarity (Schurgin et al., 2020), in line with the idea that interference is at least partially due to processing by overlapping populations of feature selective neurons in early visual cortex.

Other sensory recruitment models, however, assume that memories are maintained in a sensory-like format, but that competition occurs in higher-order areas where projections from sensory areas converge (Bouchacourt & Buschman, 2019; Swan & Wyble, 2014). For example, Bouchacourt and Buschman (2019) built a two-layer, feedforward spiking neural network where items were encoded in feature-selective sensory layers, which then sent converging random projections to a second layer, where neurons exhibited high-dimensional tuning for multiple features. Critically, inter-item interference occurs in the second layer because converging inputs from multiple sensory networks creates destructive interference when too many items are simultaneously stored. Thus, this class of model suggests that interference is *feature-general* rather than *feature-specific* (i.e. competition is only determined by overall memory load, not by inter-item similarity).

The question of feature-specific and feature-general interference has been addressed with two key lines of work about memory for conjunction objects and the mixed-category benefit. Some studies found that working memory performance in a change detection task is comparable when participants are holding in mind all features on an object compared to a single feature (Luck & Vogel, 1997), suggesting that the number of items—and not the specific visual features being stored—determines interference. However, work by Fougnie, Asplund, & Marois (2010) suggests that when high mnemonic precision was required of participants—through a continuous report task or a change detection task with high target-lure similarity—adding features to objects resulted in reduced memory precision. Fougnie & Alvarez (2011) buttressed these findings when they used a continuous report task with colorful, oriented objects and observed an independence of color and orientation report errors: one feature could be forgotten entirely, while the other was still recalled with relatively high precision. Critically, this independence was not observed for features that likely have highly overlapping neural codes, such as the length and width of objects (Fougnie & Alvarez, 2011). Taken together, these findings suggest that while there is an overall object-based benefit in visual working memory, feature-specific content nevertheless influences performance (Fougnie et al., 2013).

In addition to objects composed of simple visual features like orientation and color, prior research using real-world objects has also found mixed-category benefits that are compatible with feature-specific interference in visual working memory. Notably, Cohen et al. (2014) found that participants could remember more objects when they were from more than one category (e.g., faces and scenes) compared to when they were from one category (e.g. faces and faces). A follow-up neuroimaging experiment revealed that the size of the mixed-category benefit on a given trial was predicted by the degree of neural separability between categories (e.g., faces and scenes are processed in different neural populations; therefore, there is less cross-category competition) (Avital-Cohen & Gronau, 2021; Cohen et al., 2014; but see: Jiang et al., 2016; Mruczek et al., 2019). The mixed-category benefit overall has been replicated with simple visual features such as color, orientation, luminance, and motion (Cai et al., 2022; Gosseries et al., 2018).

The goal of the current study was to disentangle theories about the feature-generality or feature-specificity of inter-item interference during encoding and, importantly, during maintenance. In Experiment 1, we compared performance on trials with *homogeneous* displays with the same types of features (e.g., a display of colorful circles) and *heterogeneous* displays with more than one type of feature (e.g., a display of oriented bars). If inter-item interference is driven by a feature-specific component, memory precision for heterogeneous displays should be higher than memory precision for homogeneous displays because more distinct neural populations are encoding the different items. In contrast, if inter-item competition occurs in unspecialized networks that are upstream from sensory encoding, then we should observe comparable memory performance when remembering a heterogeneous display of colors and orientations and when remembering an equinumerous homogenous display of only colors or only orientations. In Experiment 2, we used retro-cues to assess whether any feature-specific interference occurred during active, online maintenance of the memoranda, or whether it could instead be explained by feature similarity during sensory encoding. Together, the studies suggest that feature similarity during encoding was the driving factor behind feature-specific interference. These results are consistent with models positing that interference in WM happens after item-specific sensory information converges in a common, more general purpose, processing mechanism (Bouchacourt & Buschman, 2019; Swan & Wyble, 2014).

Open practices statement

Experiments were preregistered on the Open Science Framework (OSF) repository (https://osf.io/h456p/). We preregistered ten experiments for this project, but for clarity and conciseness, only the most relevant experiments are reported in the manuscript body. Information about remaining experiments is available in a Supplement on OSF. Table 1 lists studies in chronological order, as well as OSF links and notes.

Table 1
Chronological Order of Experiments

Title	OSF title and link	N	Format
Experiment 1a	**Does inter-item interference occur in feature-general or feature-specific codes? (https://osf.io/tckms)**	**40**	**In-lab**
S1	Retro-cue pilot (color) (https://osf.io/vsrxc)	25	Online
S2a	Retro-cue pilot (orientation) (https://osf.io/s7qrm)	25	Online
S2b	Retro-cue pilot (orientation) 2.0 (https://osf.io/dy6nj)	25	Online
S3a	Feature interference for shapes and colors (https://osf.io/df4z9)	30	Online
S3b	Feature interference for shapes and colors (https://osf.io/5n8pm)	30	Online
Experiment 1b	**Feature interference for colors and orientations (https://osf.io/53rj4)**	**30**	**Online**
S4	Mixed-category benefit: during encoding or maintenance? (https://osf.io/3efbq)	60	Online
S5	Mixed category benefit: pre-cue edition (https://osf.io/q8shb)	60	Online
Experiment 2	**Manipulating sensory encoding and memory contents simultaneously (https://osf.io/6pbhu)**	**40**	**Online**

Experiments given in bold are included in the main manuscript body; all others can be found in the Supplement.

Experiment 1a

Experiment 1a Participants

We collected data from 44 participants from the University of California, San Diego (UCSD) community who completed the study for pay at a rate of $15/hr or for course credit. Four participants met our preregistered exclusion criteria (see below), giving us a final total of 40 participants. All were at least 18 years old, had normal or corrected-to-normal color vision, and reported no neurological disorders. All procedures were approved by UCSD's Institutional Review Board.

Experiment 1a Method

Stimuli

The stimuli and experimental procedure were programmed using MATLAB and Psychophysics Toolbox 3 (Brainard, 1997; Kleiner et al., 2007). Participants sat approximately 40 centimeters away from the computer display during the task. A chinrest was not used during the experiment, so all of the following visual angles are approximate. Stimuli were presented against a gray background with a fixation point that subtended 1 degree of visual angle. Color stimuli were circles 3° in diameter, and on each trial colors were sampled uniformly from a 360 degree CIE L*a*b color space centered at $L = 54$, $a = 18$, and $b = -8$ (Adam et al., 2017). Monitors were not calibrated to render truly equiluminant colors, but as all manipulations were within-subjects, we do not believe that this produced systematic differences between experimental conditions. Oriented bars were dark rectangles 3° in length and 1.05° in width, and angles were sampled uniformly from a 180° space.

On each trial, up to four stimuli were presented at four equidistant, fixed locations around the screen, each 6° away from the fixation point (see Figure 1). On each trial, a subset of these locations were randomly selected (depending on trial set size). Stimuli appeared for 750 ms, followed by a blank delay of

1000 ms, after which two continuous report wheels appeared at fixed locations around the entire screen. The outer wheel had an outer radius of 16°, and the inner wheel had an inner radius of 13.5°. Both wheels had an arc thickness of 2°. Whether the color or orientation wheel appeared on the outside was randomly assigned to each subject.

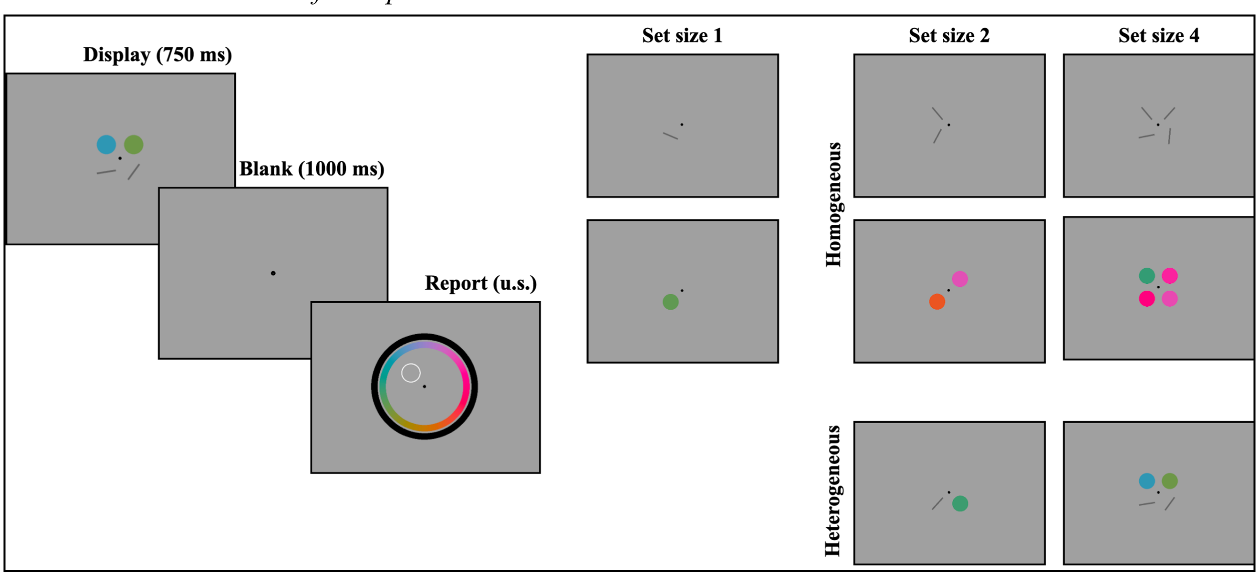
Procedure

The task (Figure 1) was a continuous report working memory task (Prinzmetal et al., 1998; Wilken & Ma, 2004; Zhang & Luck, 2008). At the start of each trial, one, two, or four items were presented on the screen. These items could be colors, oriented bars, or half colors and half oriented bars. Following the stimulus presentation and delay periods, one item from the display was probed for report by the item's location on the screen, and participants had an unlimited amount of time to make a response. Participants made a response by clicking the location on the orientation or color wheel that matched the angle or color of the probed stimulus. Despite an orientation space of 180°, the orientation wheel was a complete circle, and participants were instructed that they could click either end of the wheel.

Trial set size (one, two, or four), display condition (homogeneous, heterogeneous), and probe feature (color, orientation) were fully counterbalanced, with one small exception: set size one trials had an undefined display condition, as they are neither homogeneous nor heterogeneous. These trials were coded as "homogeneous" in the task script but were not considered homogeneous for analysis purposes. Participants completed 75 trials per condition for a total of 750 trials across the 10 conditions. These trials were spread out over 25 blocks of 30 trials each, and experimental conditions were fully counterbalanced within a block. Following each block, participants were given their average recall (in degrees), as well as the number of trials in which the feature category was incorrectly reported (e.g., participants reported an orientation when the probed stimulus was a color). Prior to the task, participants completed a set of 10 practice trials, or one trial per experimental condition, and they received feedback after each trial.

Figure 1

Procedure and Conditions for Experiment 1a



Participants saw a display of objects, followed by a delay, and then an unspeeded report period (left). We used set sizes 1, 2, and 4, and set sizes 2 and 4 could be homogeneous or heterogeneous (right)

Exclusion criteria

Participants were excluded from all analyses if more than 10 percent of trials total were feature report errors, or if any given condition had more than 20 percent feature errors (that is, reporting color when orientation was cued or vice versa). We preregistered this exclusion criteria to ensure that participants were attentive during the task and also to ensure that we obtained a sufficient number of usable trials, as we excluded all trials with feature report errors from our analyses. Previous work showing high accuracy in recalling feature categories (Awh et al., 2007; Scolari et al., 2008) suggests that these limits were not overly stringent. We also excluded a participant from all analyses if we lost more than 10% of data due to technical issues that occurred during the session (e.g., computer crashes). We preregistered that we would collect data until we had usable datasets from 40 participants. In Experiment 1a, we reached our sample size of 40 but excluded four participants who met the above criteria, so we continued data collection until we reached 40 usable datasets. In addition to those four subjects, we excluded 313 individual trials with feature report errors (1.04% of total trials).

Experiment 1a Results

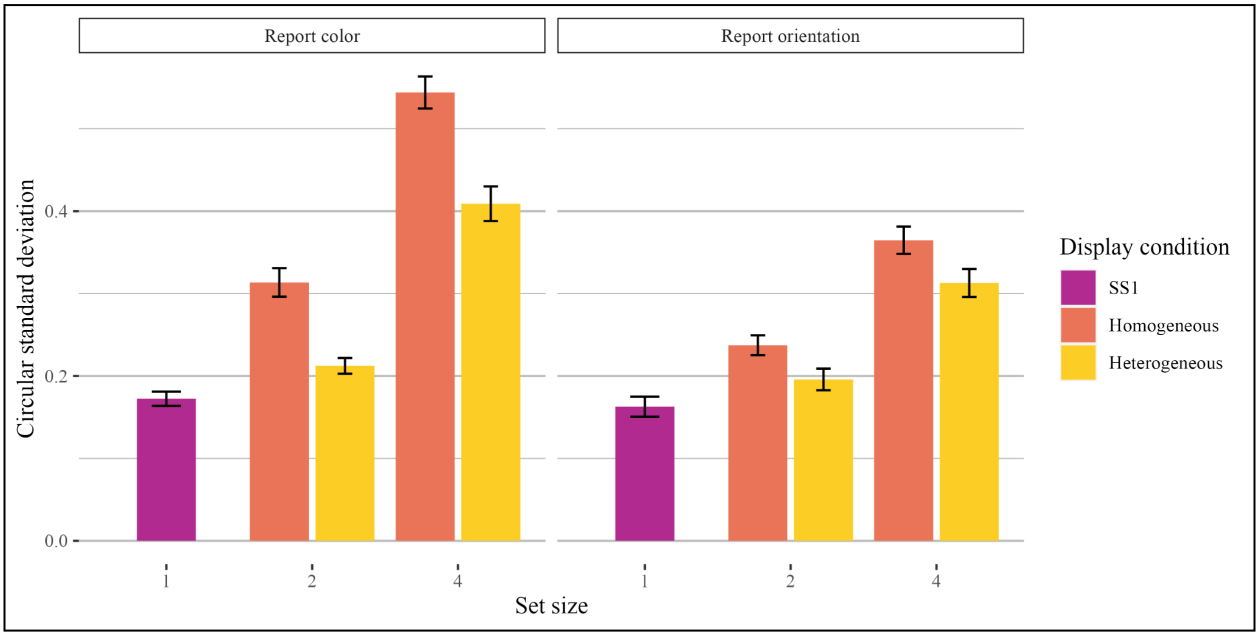
We conducted all analyses using R, version 4.3.1 (R Core Team, 2023) and *tidyverse*, version 2.0.0 (Wickham, 2023; Wickham et al., 2019). Data visualizations were created with the *ggplot2* package, version 3.4.3 (Wickham et al., 2023), as well as *viridis*, version 0.6.4 (Garnier et al., 2023).

Our primary interest was testing how heterogeneous displays affected the precision of working memory. Because color and orientation have differently-sized feature spaces (360° and 180°, respectively), comparisons were conducted separately on each probed feature. For example, we compared trials with homogeneous orientation displays and trials with heterogeneous displays where an orientation was probed for report.

Using the *circular* package, version 0.5-0 (Agostinelli & Lund, 2023), we computed the circular mean and standard deviation for each participant and experimental condition. We then ran a Bayesian two-way, repeated measures ANOVA on the set size 2 and 4 conditions using *BayesFactor,* version 0.9.12-4.4 (Morey & Rouder, 2022) and default priors. We omitted the set size 1 conditions from this analysis because these conditions have an “undefined” display condition, but we used them in follow-up planned comparisons. To assess main effects of set size and display condition, we compared a full model with set size and display condition as fixed effects to reduced models with only one or the other. Overall, Bayes Factor comparisons strongly preferred the expanded model over the model with display condition only (orientation: BF_{10} = 2.47 x 10^{29} ± 2.99%, η_p^2 = 0.70; color: BF_{10} = 2.67 x 10^{30} ± 1.75%, η_p^2 = 0.71). Set size 4 trials had lower precision (therefore, a higher circular standard deviation) than set size 2 trials. The full model with display condition was also strongly favorable (orientation: BF_{10} = 2.43 x 10^6 ± 4.24%, η_p^2 = 0.26; color: BF_{10} = 5.95 x 10^{12} ± 3.62%, η_p^2 = 0.43). Participants had lower precision in their report of homogeneous trials than heterogeneous trials. Finally, we compared a model with set size, display condition, and an interaction between the two against a reduced model without the interaction. We saw weak evidence against an interaction between set size and display condition for orientation and equivocal evidence for color (orientation: BF_{10} = 0.27 ± 4.22%; color: BF_{10} = 0.51 ± 4.56%). A plot of the mean circular standard deviations is shown in Figure 2.

Figure 2

Main Results of Experiment 1a



Results are shown separately for trials where participants reported color (left) and orientation (right). Bar plots quantify the circular standard deviation of the error distribution for each set size and display condition, and error bars represent the standard error of the mean.

Next, we conducted planned comparisons between our baseline set size 1 trials and higher set sizes. For color and orientation trials separately, we first compared the set size 1 trials to the set size 2 *homogeneous* trials. We found a main effect of set size (orientation: $BF_{10} = 6.84 \times 10^{12} \pm 1.92\%$; color: $BF_{10} = 4.11 \times 10^{9} \pm 2.98\%$). We also found a main effect of set size when we compared set size 1 and set size 2 *heterogeneous* trials (orientation: $BF_{10} = 1.22 \times 10^{5} \pm 0.72\%$; color: $BF_{10} = 4.76 \times 10^{4} \pm 1.85\%$). We also found a main effect of set size when comparing set size 1 trials to set size 4 homogeneous trials (orientation: $BF_{10} = 5.30 \times 10^{17} \pm 2.16\%$; color: $2.13 \times 10^{28} \pm 0.88\%$) and heterogeneous trials (orientation: $BF_{10} = 4.68 \times 10^{15} \pm 8.40\%$; color: $2.87 \times 10^{16} \pm 0.94\%$).

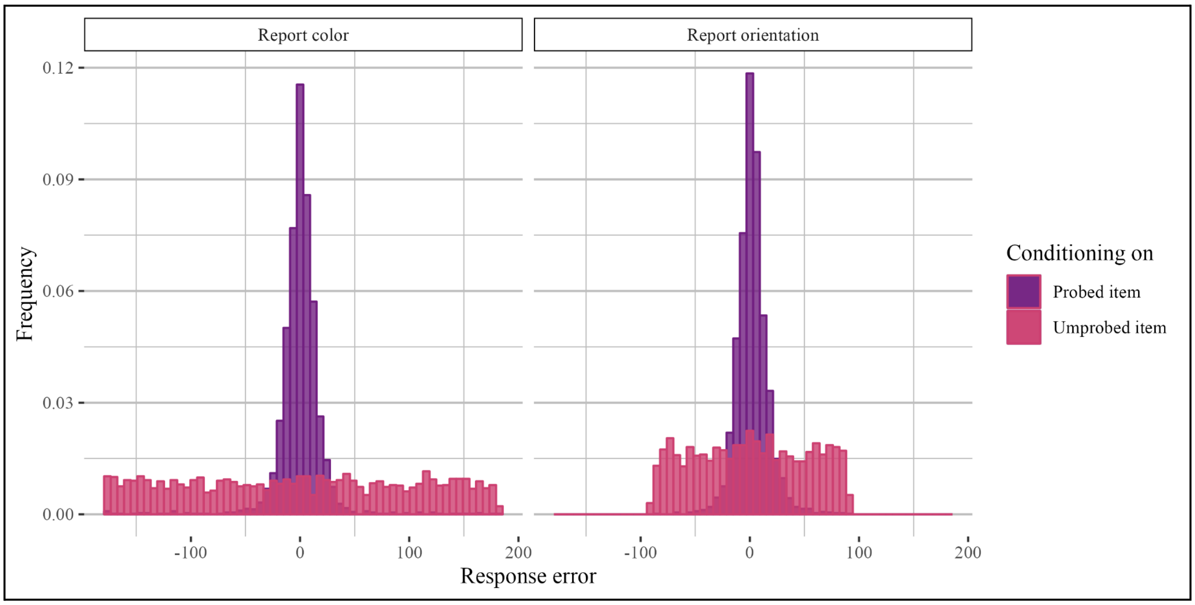
Post-hoc swap analyses

One possible explanation for the performance differences across homogeneous and heterogeneous conditions is that inter-item swapping is more likely to occur within a feature category than across a feature category (Awh et al., 2007). Displays with two colors, for example, may have lower precision than displays with one color and one orientation because participants are more likely to swap the two colors than they are the color and orientation. Further, because we excluded trials where participants clicked the wrong wheel, trials where across-category swaps occurred are not represented in this dataset. Thus, to evaluate the impact of swap errors, we took all set size 2, homogeneous trials (e.g. trials with two colors or two orientations) and computed the response error with respect to the probed item (e.g. report error) and the response error with respect to the *unprobed* item. A low response error with respect to the unprobed item is associated with a higher likelihood that the participant swapped the two items.

We used an information theoretic approach to assess uniformity of the response distribution with respect to the unprobed items (Panichello et al., 2019). Shannon Entropy is maximized for uniform distributions, so we compared the entropy of response distributions with respect to the unprobed item to the distributions of the unprobed item angles, which were drawn from a circular uniform distribution. We obtained by-participant differences in Shannon entropy for the two distributions and ran a Bayesian *t*-test to assess whether the mean difference is different from zero. The test favored the null-hypothesis of no mean difference in entropy (orientation: BF_{10} = 0.17 ± 0.05%; color: BF_{10} = 0.33 ± 0.04%). We also obtained posterior samples for the mean difference in entropy over 6000 iterations and found that the 95% posterior density interval contained zero for color and orientation reports (orientation: [-0.0167, 0.0168], color: [-0.00510, 0.0190]). These analyses suggest that the response distribution with respect to the unprobed item is relatively uniform and that swapping alone cannot explain our findings.

Figure 3

Histograms of Response Errors for Set Size 2, Homogeneous Trials (Experiment 1a)



Response error plotted with respect to the probed item and with respect to the unprobed item.

Experiment 1b

In Experiment 1b, we replicated the main finding of Experiment 1a using a web-based study and a different group of participants.

Experiment 1b Participants

We used Prolific to recruit 40 participants living in the United States. All were at least 18 years old and had normal or corrected-to-normal color vision with no colorblindness. Prior to beginning the experiment, all participants gave informed consent. All procedures were approved by UCSD’s Institutional Review Board.

Experiment 1b Method

Stimuli

We used jsPsych, version 7 (de Leeuw & Gilbert, 2023) to create the stimuli and experimental procedure, and participant data was uploaded to a secure server as a JSON file. Participants were required to complete the experiment on a desktop computer (as opposed to a smartphone or tablet), but they sat at unknown distances from the display.

Colors and orientations were chosen randomly from 360° and 180° spaces, respectively, with the constraint that colors and orientations appearing in the same trial were at least 15° apart in circular space (after Schurgin, Wixted, and Brady, (2020). Due to variation in luminance and display settings across personal computers, color stimuli may have varied across participants. While this produced a source of variance across participants, all experimental manipulations were within-subjects.

Procedure

A diagram of the trial structure is shown in Figure 4. Participants clicked a central fixation cross to begin the trial. Following each click, there was a 1500 ms delay followed by the presentation of four stimuli for 750 ms. Experimental conditions were balanced identically to set size four trials in Experiment 1a. After the offset of the stimuli, there was a 1000 ms delay, during which the placeholder circles were present but the screen was otherwise blank. At the onset of the report window, a color wheel and an orientation report wheel appeared around the placeholder circles, and the placeholder circle in the probed location had a darker border. Trials were counterbalanced so that when a heterogeneous display was shown, participants were probed to report a location with a color on half of trials and a location with an orientation on the other half. As participants moved their cursor around the report wheels, the probed location filled in with the color or orientation corresponding to their cursor's position on the wheel. Participants had unlimited time to click a location on the wheel, which locked in their response, concluding the trial. After every trial, participants were given feedback about their error in degrees, as well as feedback if they clicked the incorrect wheel.

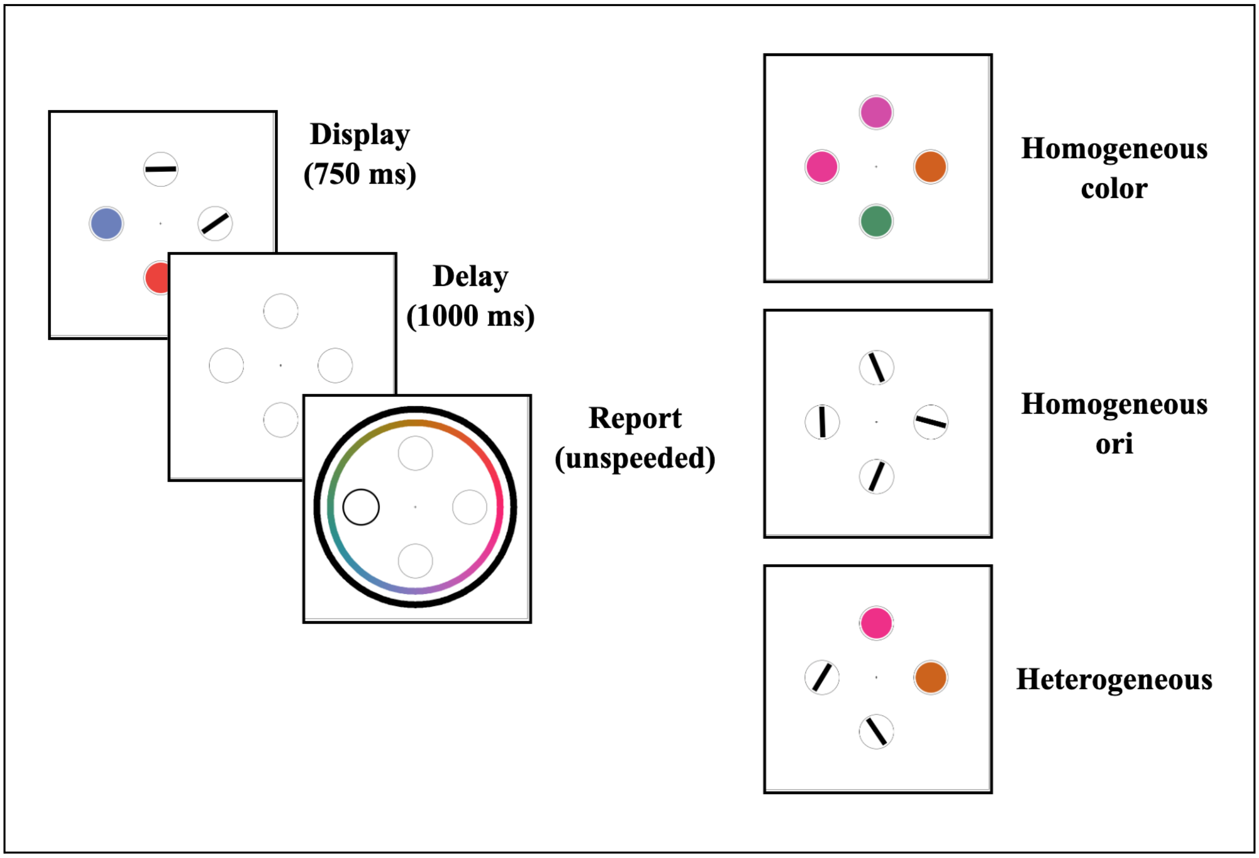
There were 20 practice trials followed by 300 main task trials, giving 75 main task trials in each of the four experimental conditions (homogeneous colors, homogeneous orientations, heterogeneous display with a color report, heterogeneous display with an orientation report).

Exclusion criteria

Participants who clicked the incorrect report wheel on more than 20% of trials in any of the four conditions were excluded from all analyses, and for all participants we excluded individual trials with an incorrect feature report. No participants were excluded, but we removed 132 individual trials where the incorrect feature wheel was clicked (1.1% of trials).

Figure 4

Procedure and Conditions for Experiment 1b



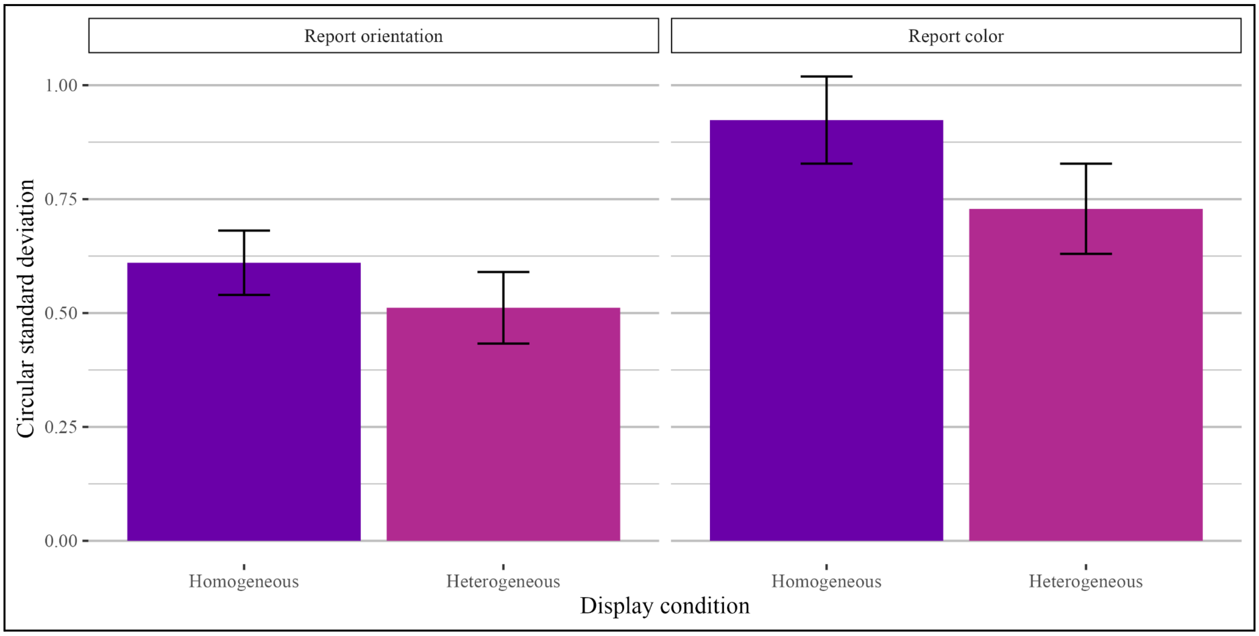
Trial structure (left) and example displays for homogeneous colors (top), homogeneous orientations (middle), and heterogeneous displays with two each of colors and orientations (bottom).

Experiment 1b Results

Performance in each condition is shown in Figure 5. We parsed JSON files using the *jsonlite* package in R, version 1.8.7 (Ooms, 2014), but data processing and aggregating methods were the same as Experiment 1a. We ran a Bayesian two-way, repeated measures ANOVA on color probe and orientation probe trials separately, with display condition (homogeneous vs. heterogeneous) as the fixed effect and participant as the random effect. Bayes Factor comparisons strongly preferred the model with display condition over the intercept-only model (orientation: BF_{10} = 4.31 x 10^3 ± 0.65%, Cohen's *d* = 0.89; color: BF_{10} = 1.59 x 10^4 ± 0.94%, Cohen's *d* = 0.99). Performance was better with heterogeneous displays.

Figure 5

Main Results for Experiment 1b



Results are shown separately for trials where participants reported orientation (left) and color (right).

Experiment 1 Discussion

In Experiment 1a, increasing set size impaired precision with both homogeneous and heterogeneous displays, in line with previous findings (Bays et al., 2009; Ma et al., 2014; Palmer, 1990). However, performance was significantly better for heterogeneous displays than homogeneous displays, suggesting at least some role of feature-specific interference. In Experiment 1b, we used a web-based study to replicate Experiment 1a at set size 4, and precision was higher for heterogeneous displays than homogeneous displays. The results add further evidence for feature-specific interference and validate the use of jsPsych and Prolific in Experiment 2.

Despite clear evidence that performance is better with heterogeneous displays, the mechanism of this benefit is unknown. While heterogeneous displays may reduce inter-item competition during maintenance, this data could also be explained by feature-similarity based competition during encoding. In Experiment 2, we used a retro-cue design (Nobre et al., 2004; for review, see Souza & Oberauer, 2016), which allowed us to manipulate the heterogeneity of the display (thereby assessing the role of similarity during encoding), as well as the heterogeneity of retro-cued items (thereby assessing the role of similarity during maintenance).

Experiment 2

Experiment 2 Participants

We used Prolific to recruit 44 participants living in the United States. Screening criteria and informed consent procedures were the same as Experiment 1b. All participants completed both sessions of the experiment.

Experiment 2 Method

Stimuli

The stimuli were identical to those used in Experiment 1b except where noted below.

Procedure

A diagram of the trial structure and experimental conditions is given in Figure 6. Clicking a central fixation cross initiated the start of the trial after a 1500 ms delay. The displays consisted of four colors on 25% of trials, four orientations on 25% of trials, and two each of colors and orientations on 50% of trials. The stimuli were present for 750 ms, followed by a 500 ms blank delay. Next, one or two of the placeholder circles had a darker border for 750 ms, indicating which item, or items, could be probed later. The retro-cue circles disappeared for 750 ms before the unspeeded report. At the onset of the report window, a color wheel and an orientation report wheel appeared, and the placeholder circle in the probed location had a darker border. Participants made only one report per trial, and the probed location was always one that was cued during the delay period.

Participants completed two sessions of equal length, and experimental conditions were counterbalanced within a session. We manipulated the display condition (homogeneous display, heterogeneous display), the feature ultimately probed for report (color, orientation), the retro-cue set size (one item, two items), and the retro-cue condition (homogeneous items retro-cued, heterogeneous items retro-cued). In total, this design produced ten experimental conditions. Experimental conditions occurred equally often over the course of the experiment, with the exception that participants completed twice as many trials with a homogeneous display of items and two items retro-cued. Although this created an imbalance in the number of trials per condition, it ensured equal numbers of trials with homogeneous and heterogeneous displays, and equal relative frequencies of retro-cue set sizes (one item cued vs. two items cued) across homogeneous and heterogeneous display conditions. Procedures for reporting were identical to Experiment 1a.

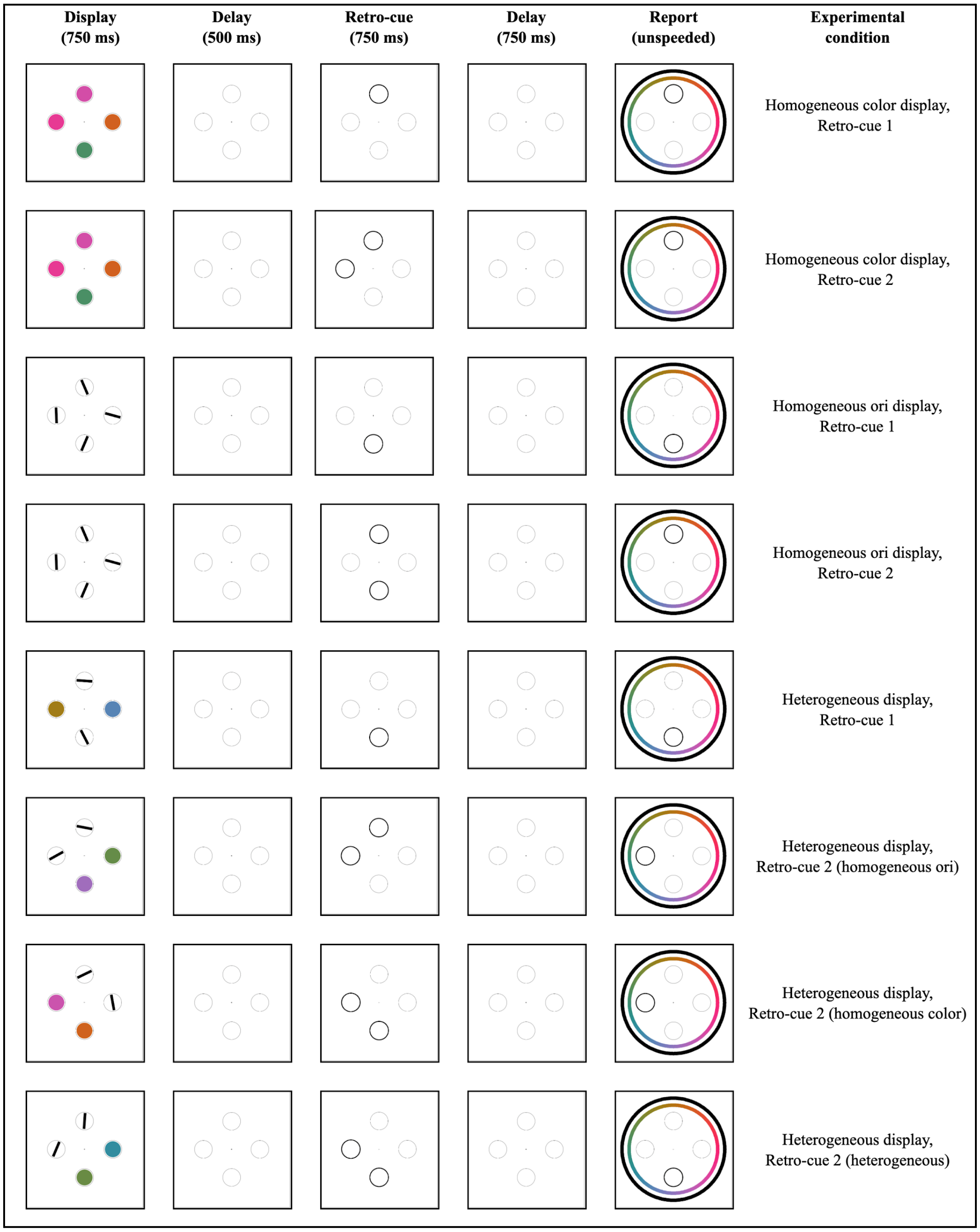
Participants completed a set of 12 practice trials, and the frequency of experimental conditions mirrored those used in the main task. There were 360 main task trials per session, giving a total of 24 practice trials and 720 main task trials. The two conditions with a homogeneous display condition and two items retro-cued had twice as many trials as other conditions, giving 120 trials in each of those two conditions and 60 trials in each of the other conditions.

Exclusion criteria and sequential data collection

We preregistered a final sample size of 40 usable participants. Because the interpretability of our experiments rests on participants using the retro-cue as intended, we preregistered a sequential data collection process to avoid wasting time and resources. The retro-cue effect is widely observed in cognitive psychology and neuroscience research (for review, see Souza & Oberauer, 2016), and the presence of a retro-cue effect in homogeneous display conditions served as a positive control. After 20 participants, we compared 1-item and 2-item retro-cue conditions for homogeneous trials and performed no additional analyses. After observing a numerical retro-cue effect for both color and orientation reports, we collected data from the remaining participants. Had we not observed a numerical effect, we would have discontinued data collection, adjusted experimental parameters, updated our preregistration, and started data collection over. Our exclusion criteria were the same as Experiment 1b. At the end of data collection, we excluded 4 participants from all analyses and 496 individual trials (or 1.7% of trials).

Figure 6

Example Displays for Experiment 2



This diagram omits conditions that differed only in the feature probed for report (e.g. heterogeneous displays where a color and orientation are retro-cued).

Experiment 2 Results

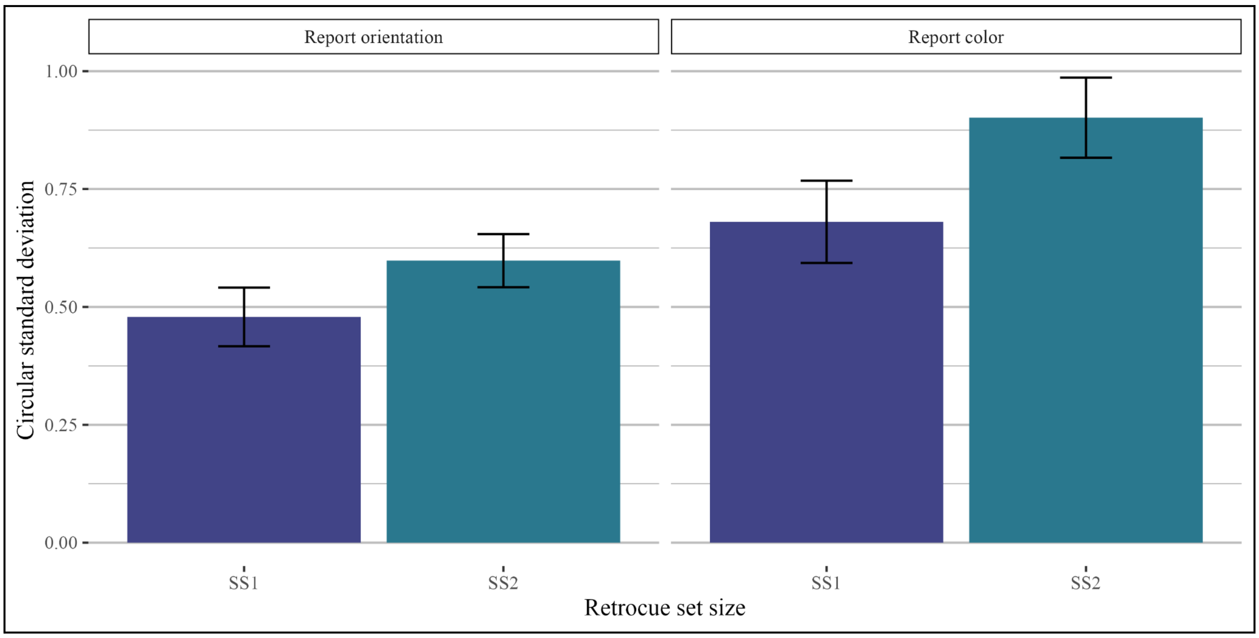
We processed and aggregated the data using the same methods as Experiment 1.

Set size

Our first comparison of interest was to look at the effect of retro-cue set size (one item versus two items retro-cued). We filtered the data to include only homogeneous display trials and ran a two-way Bayesian repeated-measures ANOVA with the retro-cue set size as a fixed effect and participant ID as a random effect. As hypothesized, model comparisons strongly preferred the full model over the intercept-only model (orientation: BF_{10} = 1.33 x 10^7 ± 2.51% , d = 1.31; color: BF_{10} = 2.39 x 10^7 ± 1.18%, d = 1.34). A plot of the circular standard deviations is shown in Figure 7. The results broadly suggest that participants were using the retro-cue as intended; therefore, when comparing trials across different retro-cue conditions, a null effect is unlikely due to non-compliance with the experiment instructions.

Figure 7

Results for Experiment 2 Homogeneous Trials



Results are shown separately for trials where participants reported color (left) and orientation (right).

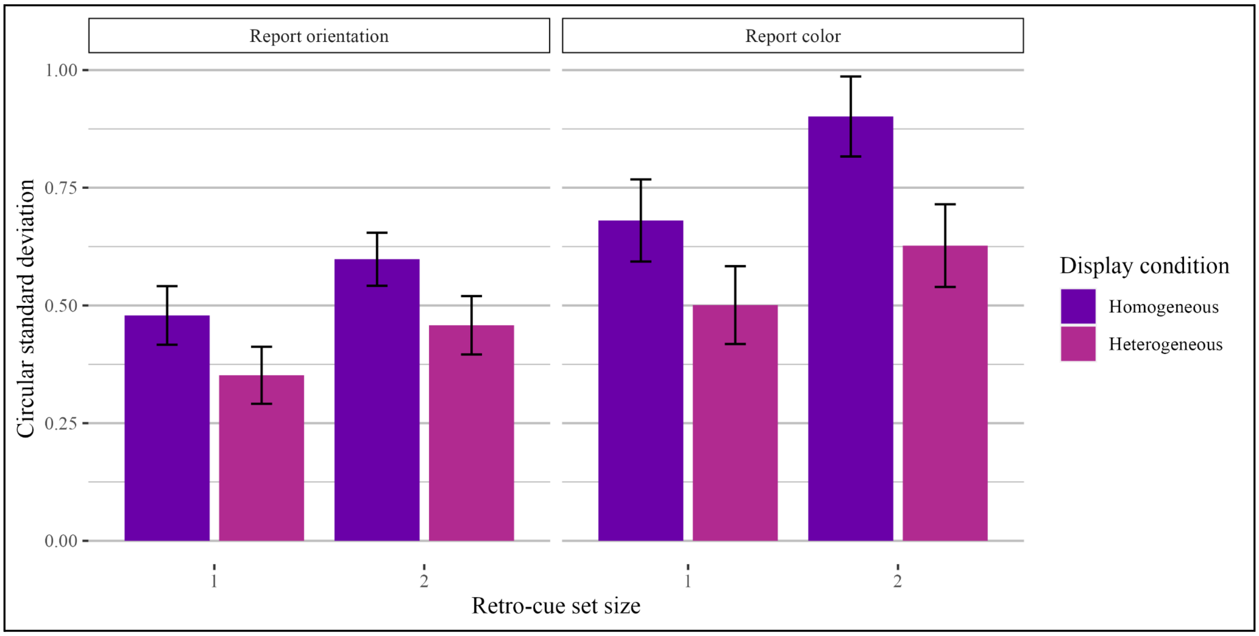
Mixed category benefit during encoding

The following analysis was mistakenly omitted from the preregistration document. To assess the mixed category benefit during encoding, we took trials with a homogeneous or SS1 *retro-cue* condition, or trials where one item was retro-cued or two items of the same feature were retro-cued. Using color trials as an example, the display condition was either four colors or two colors and two orientations, but we filtered data to include only trials with two colors retro-cued. We then compared performance across set sizes and display conditions. We ran a two-way Bayesian repeated-measures ANOVA with display condition and set size as fixed effects and participant ID as a random effect. For both color and orientation, Bayes Factor comparisons strongly favored the full model with both set size and display condition. There was a strong main effect of set size (orientation: BF_{10} = 2.76 x 10^{16} ± 2.17%, η_p^2 = 0.51; color: BF_{10} = 3.10 x 10^{11} ± 4.00%, η_p^2 = 0.40) and display condition (orientation: BF_{10} =1.27 x 10^{21} ±

1.94%, $\eta_p^2 = 0.59$; color: $BF_{10} = 4.75 \times 10^{17} \pm 4.09\%$, $\eta_p^2 = 0.53$). Performance was higher when one item was retro-cued compared to two items, and performance was higher when the display condition was heterogeneous compared to homogeneous. In other words, even when participants ultimately maintained homogeneous sets of items in WM, performance was better when the display condition was heterogeneous, replicating Experiments 1a and 1b. The results are plotted in Figures 8.

Figure 8

Results for Experiment 2 Trials with Set Size 1 or Homogeneous Retro-cue Conditions



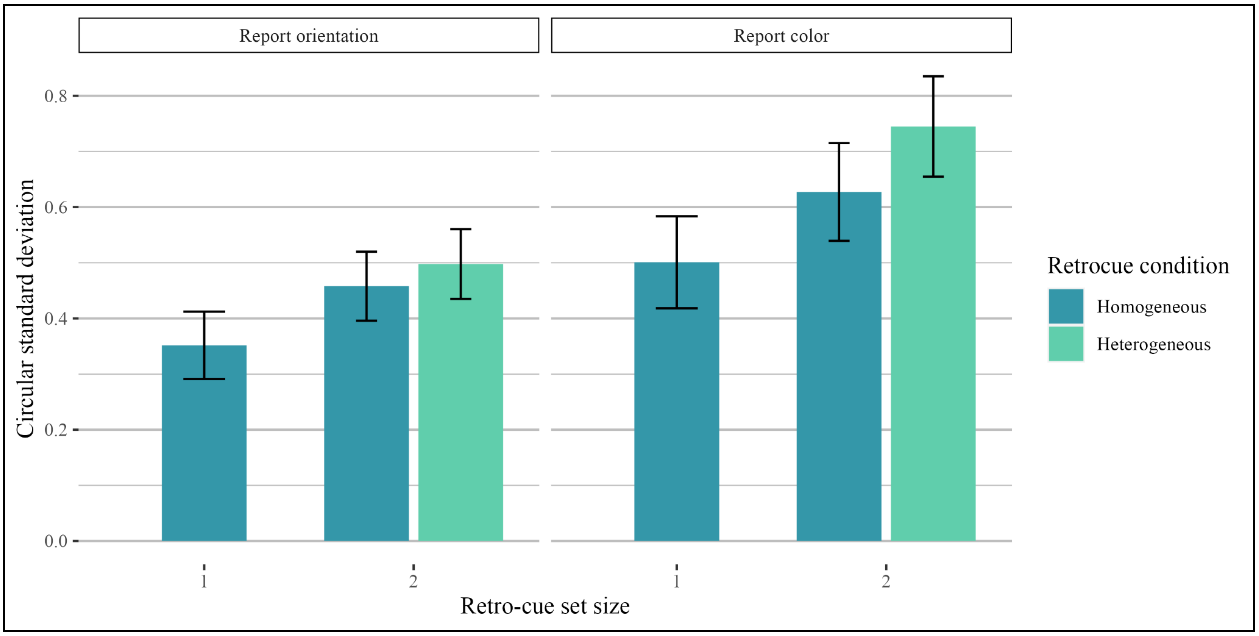
In this visualization, we kept retro-cue condition constant (retro-cue only colors, or only orientations) and visualized display condition. Results are shown separately for trials where participants reported color (left) and orientation (right).

Mixed category benefit during maintenance

Our final analysis kept display (and, thus, encoding) conditions constant and compared performance across different retro-cue conditions (see Figure 9). We analyzed only trials with a retro-cue set size of 2 and a heterogeneous display condition, and we compared trials with a homogeneous retro-cues (i.e., two of the same feature) and heterogeneous retro-cues (i.e., one color and one orientation). For color report trials, performance was better for *homogeneous* retro-cues ($BF_{10} = 2.70 \times 10^3 \pm 1.21\%$, $d = 0.84$). For orientation report trials, performance was numerically better for homogeneous retro-cues, but the Bayes Factor comparison was equivocal ($BF_{10} = 2.76 \pm 0.67\%$, $d = 0.39$). Regardless, performance differences in both color and orientation trials provided no evidence for feature-specific interference when retro-cues were involved and the properties of the stimuli during encoding were controlled.

Figure 9

Results for Experiment 2 Trials with Heterogeneous Displays



In this visualization, we kept display condition constant (two colors and two orientations) and visualized retro-cue condition. Results are shown separately for trials where participants reported color (left) and orientation (right).

Experiment 2 Discussion

In Experiment 2, we observed feature-specific interference during encoding, consistent with Experiments 1a and 1b. However, the previously-observed performance benefits for heterogeneous sets of items disappeared when these items were retro-cued. In other words, when participants encoded a heterogeneous display, and we compared performance when two colors or two orientations were retro-cued or one of each feature was retro-cued, performance was either equivocal between the conditions or better for two colors. Further, the null findings were unlikely a result of non-compliance with experiment instructions, as performance was better when one item was retro-cued than when two-items were retro-cued.

General Discussion

The goal of the present work was to test different accounts of WM that propose feature-specific or feature-general sources of competition and inter-item interference. In Experiment 1a, increasing the display set size produced a cost in mnemonic precision regardless of whether the displays were homogeneous or heterogeneous. However, for set sizes two and four, mnemonic precision was better for heterogeneous displays compared to homogeneous displays. These findings replicate and extend previous research on the mixed-category benefit (Avital-Cohen & Gronau, 2021; Cohen et al., 2014). More importantly, these results suggest that inter-item competition occurs in both a feature-general manner as more items are remembered, and in a feature-specific manner that depends on item similarity. In Experiment 1b, we replicated the findings of Experiment 1a at set size four and validated the use of online experiments for these studies more generally. Experiment 2 used a retro-cueing design and revealed that encoding a display of heterogeneous items is advantageous for mnemonic precision but that once a given set of items are encoded into working memory, the feature-specific interference disappears and there is no benefit associated with remembering heterogeneous sets (with even mild evidence that homogeneous displays are remembered with higher precision). Thus, feature-specific interference likely arises during sensory encoding, but once encoded, there is no evidence for feature-specific competition during maintenance in WM.

One key motivation for our experiment is that different instantiations of sensory recruitment models of working memory make qualitatively different predictions about the role of feature similarity in inter-item interference. For example, some models assume that inter-item competition occurs via competition in higher-order processing stages that aggregate information from many feature-selective sensory neurons tuned to different features in earlier processing stages (Bouchacourt & Buschman, 2019; Swan & Wyble, 2014). In terms of behavior, increasing the set size should reduce working memory performance because of more convergent input to high-order areas, but the combinations of feature types should not matter. Overall, the results of Experiment 2 indicate that once items are encoded into working memory, the nature of the inputs matters little—a finding consistent with this class of model. Bouchacourt and Buschman (2019) even made the prediction that increasing inter-item similarity may increase the stability of representations in working memory, consistent with prior work and our observation that, once encoded, homogeneous sets of items were remembered with slightly higher precision (Kiyonaga et al., 2017; Lin & Luck, 2009). Thus, our results dovetail with this class of model, particularly with versions of the architecture that allow for competition during sensory encoding via lateral excitatory connections between like-tuned units in the sensory layers, which could account for the lower performance we observed when *encoding* homogeneous displays.

The finding that memory representations are robust to feature-specific interference suggests a prominent role of higher-order regions in mediating inter-item competition. However, our data are agnostic about whether item-specific information is stored in a sensory-like code or whether it is re-coded into a non-sensory format and stored in higher-order brain areas (Christophel et al., 2012; Ester et al., 2016; Iamshchinina et al., 2021; Serences, 2016; Xu, 2017, 2020). For example, Bouchacourt & Buschman (2019) proposed that inter-item interference originates due to destructive interference in higher layers where units receive convergent inputs from many sensory neurons with different feature-specific tuning functions (Swan & Wyble, 2014). However, the disruption of memory representations is realized via the backpropagation of signals from higher layers to the sensory layers where information about each remembered item is actually maintained. Thus our observation that competition does not have a strong feature-selective component is consistent with prior work demonstrating that high-fidelity mnemonic information are maintained in sensory cortices (Christophel et al., 2012; Emrich et al., 2013; Favila et al., 2020; Harrison & Tong, 2009; Rademaker et al., 2019; Serences et al., 2009). Equally, our results could be accommodated by models in which sensory regions are active during encoding, but activity in higher-order areas forms the basis for maintaining active memory representations and behavioral read-out (Bettencourt & Xu, 2016; Xu, 2017). The behavioral data presented here cannot adjudicate between these two models of storage without further constraints provided by neural data.

In sum, our data suggest that inter-item interference is feature-specific during sensory encoding but feature-general once items are in working memory. These results are consistent with theoretical accounts of WM in which populations of unspecialized neurons in higher-order brain regions aggregate information from sensory-tuned neural populations early in visual processing. More broadly, we provide empirical support for the hypothesis that coordinated communication between highly-specialized and highly-flexible neurons gives rise to working memory's flexible and adaptive nature.

References

Adam, K. C. S., Rademaker, R. L., & Serences, J. T. (2022). Evidence for, and challenges to, sensory recruitment models of visual working memory. In *Visual Memory* (pp. 5–25). Routledge.

Adam, K. C. S., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology*, *97*, 79–97.

Agostinelli, C., & Lund, U. (2023). *R package circular: Circular Statistics (version 0.5-0)*. https://CRAN.R-project.org/package=circular

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.

Avital-Cohen, R., & Gronau, N. (2021). The asymmetric mixed-category advantage in visual working memory: An attentional, not perceptual (face-specific) account. *Journal of Experimental Psychology. Human Perception and Performance*, *47*(6), 852–868.

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, *18*(7), 622–628.

Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(10), 3632–3645.

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.

Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience*, *19*(1), 150–157.

Bouchacourt, F., & Buschman, T. J. (2019). A Flexible Model of Working Memory. *Neuron*, *103*(1), 147–160.e8.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 4.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Cai, Y., Fulvio, J. M., Samaha, J., & Postle, B. R. (2022). Context Binding in Visual Working Memory Is Reflected in Bilateral Event-Related Potentials, But Not in Contralateral Delay Activity. *eNeuro*, *9*(6). https://doi.org/10.1523/ENEURO.0207-22.2022

Christophel, T. B., Hebart, M. N., & Haynes, J.-D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(38), 12983–12989.

Chunharas, C., Rademaker, R. L., Brady, T. F., & Serences, J. T. (2022). An adaptive perspective on visual working memory distortions. *Journal of Experimental Psychology. General*. https://doi.org/10.1037/xge0001191

Cohen, M. A., Konkle, T., Rhee, J. Y., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(24), 8955–8960.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87–114; discussion 114–185.

de Leeuw, J. R., & Gilbert, R. A. (2023). jsPsych: Enabling an Open-Source Collaborative Ecosystem of Behavioral Experiments. *Journal of Open Source*. https://joss.theoj.org/papers/10.21105/joss.05351.pdf

D'Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory. *Annual Review* *of Psychology*, *66*(1), 115–142.

Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed Patterns of Activity in Sensory Cortex Reflect the Precision of Multiple Items Maintained in Visual Short-Term Memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(15), 6516–6523.

Ester, E. F., Anderson, D. E., Serences, J. T., & Awh, E. (2013). A neural measure of precision in visual working memory. *Journal of Cognitive Neuroscience*, *25*(5), 754–761.

Ester, E. F., Rademaker, R. L., & Sprague, T. C. (2016). How Do Visual and Parietal Cortex Contribute to Visual Short-Term Memory? [Review of *How Do Visual and Parietal Cortex Contribute to Visual Short-Term Memory?*]. *eNeuro*, *3*(2). https://doi.org/10.1523/ENEURO.0041-16.2016

Favila, S. E., Lee, H., & Kuhl, B. A. (2020). Transforming the Concept of Memory Reactivation. *Trends in Neurosciences*, *43*(12), 939–950.

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). https://doi.org/10.1167/11.12.3

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision*, *10*(12), 27–27.

Fougnie, D., Cormiea, S. M., & Alvarez, G. A. (2013). Object-based benefits without object-based representations. *Journal of Experimental Psychology. General*, *142*(3), 621–626.

Garnier, Simon, Ross, Noam, Rudis, Robert, Camargo, Pedro, A., Sciaini, Marco, Scherer, & Cédric. (2023). *viridis(Lite) - Colorblind-Friendly Color Maps for R*. https://doi.org/10.5281/zenodo.4679423

Gayet, S., Paffen, C. L. E., & Van der Stigchel, S. (2018). Visual Working Memory Storage Recruits Sensory Processing Areas. *Trends in Cognitive Sciences*, *22*(3), 189–190.

Gold, J. M., & Luck, S. J. (2023). Working Memory in People with Schizophrenia. *Current Topics in Behavioral Neurosciences*, *63*, 137–152.

Golomb, J. D. (2015). Divided spatial attention and feature-mixing errors. *Attention, Perception & Psychophysics*, *77*(8), 2562–2569.

Gosseries, O., Yu, Q., LaRocque, J. J., Starrett, M. J., Rose, N. S., Cowan, N., & Postle, B. R. (2018). Parietal-Occipital Interactions Underlying Control- and Representation-Related Processes in Working Memory for Nonspatial Visual Features. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *38*(18), 4357–4366.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632–635.

Iamshchinina, P., Christophel, T. B., Gayet, S., & Rademaker, R. L. (2021). Essential considerations for exploring visual working memory storage in the human brain. *Visual Cognition*, *29*(7), 425–436.

Jiang, Y. V., Remington, R. W., Asaad, A., Lee, H. J., & Mikkalson, T. C. (2016). Remembering faces and scenes: The mixed-category advantage in visual working memory. *Journal of Experimental Psychology. Human Perception and Performance*, *42*(9), 1399–1411.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, *139*(1), 181–193.

Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial Dependence across Perception, Attention, and Memory. *Trends in Cognitive Sciences*, *21*(7), 493–497.

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?*

https://pure.mpg.de/rest/items/item_1790332/component/file_3136265/content

Lewis-Peacock, J. A., & Norman, K. A. (2014). Competition between items in working memory leads to forgetting. *Nature Communications*, *5*(1), 5768.

Lin, P.-H., & Luck, S. J. (2009). The Influence of Similarity on Visual Working Memory Representations. *Visual Cognition*, *17*(3), 356–372.

Liu, K., & Jiang, Y. (2005). Visual working memory for briefly presented scenes. *Journal of Vision*, *5*(7), 650–658.

Lively, Z., Robinson, M. M., & Benjamin, A. S. (2021). Memory fidelity reveals qualitative changes in interactions between items in visual working memory. *Psychological Science*, *32*(9), 1426–1441.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356.

Morey, R. D., & Rouder, J. N. (2022). *BayesFactor: Computation of Bayes Factors for Common Designs*. https://richarddmorey.github.io/BayesFactor/

Mruczek, R. E. B., Killebrew, K. W., & Berryhill, M. E. (2019). Individual differences reveal limited mixed-category effects during a visual working memory task. *Neuropsychologia*, *122*, 1–10.

Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., & Mesulam, M. M. (2004). Orienting attention to locations in perceptual versus mental representations. *Journal of Cognitive Neuroscience*, *16*(3), 363–373.

Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review*, *124*(1), 21–59.

Ooms, J. (2014). The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. In *arXiv:1403. 2805 [stat. CO]*. https://arxiv.org/abs/1403.2805

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology. Human Perception and Performance*, *16*(2), 332–350.

Panichello, M. F., DePasquale, B., Pillow, J. W., & Buschman, T. J. (2019). Error-correcting dynamics in visual working memory. *Nature Communications*, *10*(1), 3366.

Pratte, M. S., & Tong, F. (2014). Spatial specificity of working memory representations in the early visual cortex. *Journal of Vision*, *14*(3), 22.

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology. Human Perception and Performance*, *24*(1), 261–282.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. *Nature Neuroscience*, *22*(8), 1336.

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schecklmann, M., Dresler, T., Beck, S., Jay, J. T., Febres, R., Haeusler, J., Jarczok, T. A., Reif, A., Plichta, M. M., Ehlis, A.-C., & Fallgatter, A. J. (2011). Reduced prefrontal oxygenation during object and spatial visual working memory in unpolar and bipolar depression. *Psychiatry Research*, *194*(3), 378–384.

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Publisher Correction: Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-020-00993-7

Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic Bulletin & Review*, *15*(1), 215–222.

Scotti, P. S., Hong, Y., Leber, A. B., & Golomb, J. D. (2021). Visual working memory items drift apart due to active, not passive, maintenance. *Journal of Experimental Psychology. General*, *150*(12), 2506–2524.

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, *128*, 53–67.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-Specific Delay Activity in Human Primary Visual Cortex. *Psychological Science*, *20*(2), 207–214.

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*(7), 1839–1860.

Sprague, T. C., Ester, E. F., & Serences, J. T. (2014). Reconstructions of Information in Visual Spatial Working Memory Degrade with Memory Load. *Current Biology: CB*, *24*(18), 2174–2180.

Swan, G., & Wyble, B. (2014). The binding pool: a model of shared neural resources for distinct items in visual working memory. *Attention, Perception & Psychophysics*, *76*(7), 2136–2157.

Wickham, H. (2023). *tidyverse: Easily Install and Load the Tidyverse*. https://tidyverse.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D. (2023). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. https://ggplot2.tidyverse.org

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 1120–1135.

Wyble, B., & Swan, G. (2015). Mapping the spatiotemporal dynamics of interference between two visual targets. *Attention, Perception & Psychophysics*, *77*(7), 2331–2343.

Xu, Y. (2017). Reevaluating the Sensory Account of Visual Working Memory Storage. *Trends in Cognitive Sciences*, *21*(10), 794–815.

Xu, Y. (2020). Revisit once more the sensory storage account of visual working memory. *Visual Cognition*, *28*(5-8), 433–446.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235.