THE UNIVERSITY OF CHICAGO

STORAGE IN VISUAL WORKING MEMORY RECRUITS A CONTENT-INDEPENDENT POINTER SYSTEM

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

WILLIAM SOMERSET THYER

CHICAGO, ILLINOIS

AUGUST 2023

Content

List of Figures iv

List of Tables vi

Acknowledgements vii

1. Storage in visual working memory recruits a content-independent pointer system 1
2. Stronger evidence for feature-independent pointers in visual working memory 30
3. Spatial attention and working memory gating are distinct forms of voluntary attentional control 51

Bibliography 73

List of Figures

Figure 1. Task schematics for Experiments 1, 2, & 3 8

Figure 2. Change detection accuracy for each set size in each experiment 14

Figure 3. Classification accuracy over time for each experiment 16

Figure 4. Classification accuracy of single-feature load for set sizes 1 vs. 2, 2 vs. 3, and 3 vs. 4. 16

Figure 5. Individual differences in classification accuracy correlated with working memory capacity. 17

Figure 6. Accuracy for single-feature load classification. 18

Figure 7. Accuracy for classification trained on single-feature items and tested on conjunction items. 19

Figure 8. Classifier output (trained on single-feature items) compared to the predicted output of the feature load and pointer hypotheses. 21

Figure 9. Examples of set size 4 grouped and ungrouped memory arrays. 22

Figure 10. Distance from the classification hyperplane for set size 2 ungrouped, set size 4 grouped, and set size 4 ungrouped trials across time. 24

Figure 11. Task schematics for an examples set size 3 trial in the whole-field change-detection used in both experiments. 35

Figure 12. Change detection accuracy for each set size in both experiments. 40

Figure 13. Event-related potentials for each condition in both experiments. 41

Figure 14. Load classification accuracy for each attended feature over time for both experiments. 42

Figure 15. Accuracy for within- and across-feature load classification. 45

Figure 16. Accuracy for attended feature classification in both experiments. 46

Figure 17. Accuracy for motion coherence level classification. 47

Figure 18. Task schematics for each condition and example set size 1 trials for each experiment. 55

Figure 19. Task accuracy for each condition in both experiments. 61

Figure 20. Hyperplane and CTF results for Experiment 1. 63

Figure 21. Hyperplane and CTF results for Experiment 2. 65

List of Tables

Table 1. Tukey HSD analysis results for distance from the hyperplane at each time window in Experiment 1. 64

Table 2. Tukey HSD analysis results for distance from the hyperplane at each time window in Experiment 2. 68

Acknowledgments

Thanks to Ed Awh for his mentorship.

Thanks to Ed Vogel, Monica Rosenberg, and Marc Berman for their guidance and feedback.

Thanks to Colin Quirk, Megan deBettencourt, William Ngiam, and Nicole Hakim for taking so much time to teach me new things.

Thanks to my parents for fostering a sense of curiosity.

Chapter 1: Storage in visual working memory recruits a content-independent pointer system

Introduction

A central goal of cognitive neuroscience has been to understand the neural underpinnings of working memory (WM), an online memory system that is thought to be critical for virtually all forms of intelligent behavior. Significant progress has been made by focusing on stimulus-specific neural activity that tracks the features of the items stored in working memory. In both animal and human subjects, WM storage has been shown to elicit sustained activity in neural units or brain regions that are selective for the particular items held in mind (D'Esposito & Postle, 2015; Funahashi et al., 1989; Fuster & Jervey, 1981; Goldman-Rakic, 1995; Harrison & Tong, 2009; Rademaker et al., 2019; Serences et al., 2009). The motivation for these studies is clear, as they have the potential to elucidate the memory engrams (Poo et al., 2016) that allow us to hold specific ideas in mind.

Nevertheless, a distinct category of studies has focused instead on neural signals that track the *number* of items stored in working memory, rather than the content of those representations (Adam et al., 2020; Todd & Marois, 2004; Vogel & Machizawa, 2004; Xu & Chun, 2006). For example, Vogel and Machizawa (2004) used scalp electroencephalogram (EEG) recordings to observe a sustained negative slow wave in posterior electrodes contralateral to the items stored in working memory. This contralateral delay activity (CDA) persists throughout the delay period, reaches a

plateau when behavioral estimates of memory capacity are exceeded, and is a robust predictor of individual differences in the capacity of visual working memory (Luria et al., 2016). This kind of load-sensitive neural measure has provided insight into how observers control access to this limited online workspace (McNab et al., 2008; Vogel et al., 2005), the role of working memory in complex tasks such as multiple object tracking (Drew & Vogel, 2008) and visual search (Carlisle et al., 2011; Gunseli et al., 2014), and the relationship between working memory capacity and other cognitive abilities (Unsworth et al., 2014, 2015).

Although it is clear that load-sensitive neural signals have been potent tools for studying working memory, important questions remain regarding the computational role of this class of neural activity. Given past evidence for sustained stimulus-specific neural activity during WM storage, one possibility is that load-sensitive signals index the feature-selective neural activity required for storage. Here, however, we present evidence for neural activity that indexes a qualitatively different cognitive operation from the representation of content, per se. There has been longstanding interest in the cognitive operations that support object individuation -- the segmentation of objects from the background and from other objects -- and the binding of an item's features into an integrated percept that can be tracked in a dynamic visual scene. Kahneman et al. (1992) proposed the "object file" as a mechanism for registering specific tokens in the visual field to support the continuous tracking of those items through time and space (Kahneman et al., 1992). Likewise, Pylyshyn (2009) described "Fingers of Instantiation" (FINSTs) as a mechanism for indexing visual tokens, thereby enabling perception to

unfold over time despite changes in appearance or spatial position (Z. W. Pylyshyn, 2009). Thus, both theories describe a kind of spatiotemporal “pointer” system that supports the apprehension and tracking of individuated items, while the stored *content* about each item in memory is maintained via parallel but distinct while distinct mechanisms support the maintenance of each item’s attended features.

Our hypothesis is that load-sensitive neural signals reflect the deployment of these spatiotemporal pointers. Although the pointer construct was developed in the context of attentional tracking tasks, WM storage can also be construed as the sustained deployment of attention towards *internal* representations (Awh & Jonides, 2001). Indeed, multiple models of visual WM have embraced the idea of separable neural processes for the storage of content on the one hand, and the individuation and binding of those representations on the other (Balaban et al., 2019; Bouchacourt & Buschman, 2019; Oberauer, 2019; Swan & Wyble, 2014; Xu & Chun, 2009). For example, Swan and Wyble (2014) postulate a neural “binding pool” that serves to link together the multiple features of stored items, supporting their representation as individuated tokens. Likewise, Xu and Chun (2009) argued that object individuation and object identification are realized in independent stages of processing, with distinct cortical regions supporting each function (Xu & Chun, 2006). Thus, there is clear motivation to postulate the existence of load-sensitive neural signals that index a content-independent aspect of working memory. Our primary conclusion is that EEG activity measured during WM storage provides evidence of precisely this kind of neural operation.

We used a recently developed multivariate approach that uses the scalp topography of EEG activity to decode the number of individuated items held in visual working memory (Adam et al., 2020). Although past work has found univariate signals that index the number of items stored in working memory, there are several reasons why multivariate load detection (mvLoad) provides a more powerful testbed for characterizing the properties of load-sensitive neural activity. First, mvLoad is far more sensitive, enabling above-chance tracking of the number of items stored even with single trials of EEG activity. Second, mvLoad analyses reveal a multivariate signature of WM storage that generalizes from the trained dataset to novel human observers, and across significant variations in task design (e.g., lateralized versus whole field memory displays); thus the method is able to isolate load-sensitive activity more decisively than prior approaches. Finally, mvLoad accuracy robustly predicts individual differences in WM capacity, showing that it taps into an integral aspect of this online memory system.

We focused on three clear predictions for the properties of load-sensitive neural activity that is separable from the maintenance of specific visual details. First, the activity should precisely track the number of individuated representations that are encoded into memory, independent of variations in stimulus-driven activity. Second, the activity should generate a load signature that *generalizes* across the storage of distinct classes of visual information. Third, that signature should generalize across strong variations in the amount of *information* stored about each item, establishing that it tracks the number of individuated representations rather than the total amount of information stored. To anticipate the results, three studies using the mvLoad analytic approach confirm all of

these predictions, thereby providing critical new evidence for theories of WM capacity that distinguish between the storage of featural details, and the indexing of individuated items within visual working memory. We propose that this content-independent signature of WM load indexes the deployment of spatiotemporal pointers (e.g., Pylyshyn, 2009; Khaneman et al., 1992) that enable the individuation, binding, and monitoring of attended objects.

Methods

Subjects

Experiments 1-3 included 95 separate data collection sessions (42 in Experiment 1, 33 in Experiment 2, and 21 in Experiment 3), with 50 unique volunteers participating for monetary compensation ($15/hr). A total of 20 volunteers participated in all 3 studies, allowing us to implement cross-training analyses across experiments. For participants who completed multiple experiments, each experiment was done in a separate EEG session. Subjects were between 18 and 35 years old, reported normal or corrected-to-normal visual acuity, and provided informed consent according to procedures approved by the University of Chicago Institutional Review Board.

Experiment 1

Our target sample was 30 subjects in Experiment 1. 42 volunteers participated in Experiment 1 (25 female; mean age = 23.8, SD = 4.5). 9 subjects were excluded from the final sample for the following reasons: we were unable to prepare the subject for EEG (n = 2); the subject did not complete enough blocks of the task (n = 5); the

subject's data was unintentionally overwritten (n = 1), and too many trials were rejected due to eye movements (see Eye Movement Controls, n = 1). The final sample size was 33 (20 female; mean age = 24.33 years, SD = 4.76). We overshot our target sample size by 3 because we needed enough subjects to complete all 3 experiments and some could not return.

Experiment 2

Our target sample was 30 subjects in Experiment 1. 33 volunteers participated in Experiment 1 (18 female; mean age = 25.39, SD = 4.30). 2 subjects were excluded from the final sample for the following reasons: the subject did not complete enough blocks of the task (n = 2). The final sample size was 31 (18 female; mean age = 25.32 years, SD = 4.07). We overshot our target sample size by 1 because we needed enough subjects to complete all 3 experiments and some could not return.

Experiment 3

Our target sample was 20 subjects in Experiment 3. 20 volunteers participated in Experiment 3 (13 female; mean age = 25.45, SD = 4.07). No subjects were excluded from the final sample.

Apparatus

We tested the subjects in a dimly lit, electrically shielded chamber. Stimuli were generated using PsychoPy. Subjects viewed the stimuli on a gamma-corrected 24" LCD

monitor (refresh rate: 120 Hz, resolution 1080 x 1920 pixels) with their chin on a padded chin rest at a viewing distance of 75 cm.

Luminance-balanced displays

Stimuli were presented against a mid-gray background (~61 cd/m 2). Memory arrays included 1-4 to-be-remembered items. Ignored placeholder items also appeared in the memory array such that each array had a total of 5 items. The placeholder items were a gray (RGB-value: 166,166,166) that matched the average luminance of all possible colors in the color set.

Task Procedures

All three experiments used a whole-field change detection task. On each trial, a memory array appeared containing five total items. There were 1-4 colored items to be remembered, and the remainder of items were gray placeholder items to balance area and luminance across set size conditions (see Luminance-balanced displays for more detail). Memory and placeholder items were positioned with 1 item per quadrant, plus the 5th item which was placed in a randomly selected quadrant. 2 memory items never appeared in 1 quadrant together. All items were placed at least 4° apart. Participants viewed a memory array (250 ms), remembered the items across a delay (1000 ms), were probed on one item, and reported whether the probed item was the same as or different from the remembered item (unspeeded). Participants completed 14 blocks of 120 trials, for a total of 1680 trials per session (420 per set size). 2 subjects only

completed 1,348 and 1,440 trials each. EEG acquisition duration was between 73 and 132 minutes, with an average of 105 minutes.

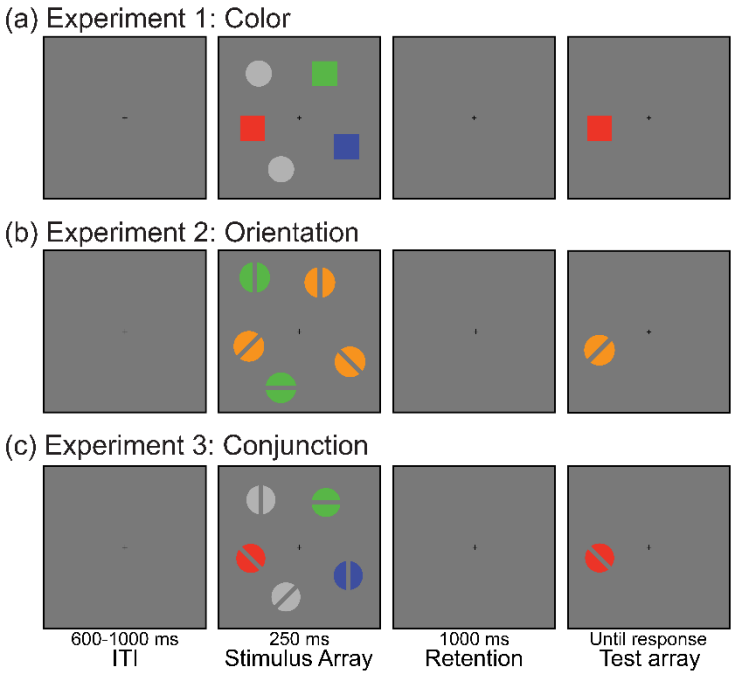

Figure 1 Task schematics for an example set size 3 trial in the whole-field change detection task used in all three experiments. (a) In Experiment 1. participants remembered the colored squares while ignoring the grey placeholders. (b) In Experiment 2. at the start of each blook, a color cue informed participants to attend and remember either the orange or the green orientations while ignoring the unced color. (c) In Experiment 3. participants remembered both the color and the orientation of each item. During change trials, one of the features (randomly selected) in the test item would change.

Experiment 1: Color

In Experiment 1, the memory items were colored squares (width = 2°) (Figure 1a). The colors were randomly sampled without replacement from a set of 7 colors (red = 255, 0, 0; green = 0, 255, 0; blue = 0, 0, 255, yellow = 255, 255, 0; purple = 255, 0, 255; teal = 0, 255, 255; orange = 255, 128, 0). Circular gray placeholders (radius = 1.13°) of the same area as the memory items also appeared during the memory array such that each display contained 5 total objects. One potential concern is that the spatial frequency of displays covaried with load because the colored squares had a higher spatial frequency than the circular placeholders. Although this raises a possible alternative explanation of load decoding in the color condition, there was no similar concern with the displays used in Experiments 2 and 3.

Experiment 2: Orientation

In Experiment 2, the memory items were circles (radius = 1.3°) with oriented bars cut out of the middle (height = 2.6°, width = .5°) such that they were the same area as the items in Experiment 1 (Figure 1b). The possible orientations were 0, 90, 180, and 270 degrees and they were sampled without replacement for each trial. The placeholder items were the same shape. Each block, either orange or green was indicated as the target color for that block. Subjects were instructed to remember the orientation of the stimuli presented in the target color, and to ignore the stimuli presented in the other color. Both the orange and green were luminance matched to the average luminance of the color set in Experiment 1 (orange = 255, 155, 55; green = 75, 208, 75). Thus, luminance was perfectly balanced across set size conditions. For example, a trial that contains 1 orange and 4 green items would be a set size 1 trial in a “target orange” block but a set size 4 trial in a “target green” block.

Experiment 3: Conjunction

In Experiment 3, each memory item included both an orientation and a color and was the same shape and size as items in Experiment 2. Both features were independently sampled without replacement from the same color and orientation values used in Experiments 1 and 2 (Figure 1c). The placeholders were the luminance-matched grey from Experiment 1. In change trials, only one attribute (color or orientation) changed, with color and orientation changes occurring equally often.

EEG acquisition

We recorded EEG activity from 30 active Ag/AgCl electrodes mounted in an elastic cap (Brain Products actiCHamp, Munich, Germany). We recorded from International 10-20 sites: Fp1, Fp2, F7, F3, Fz, F4, F8, FT9, FC5, FC1, FC2, FC6, FT10, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, O2. Two additional electrodes were affixed with stickers to the left and right mastoids, and a ground electrode was placed in the elastic cap at position Fpz. All sites were recorded with a right-mastoid reference and were re-referenced offline to the algebraic average of the left and right mastoids. We recorded electrooculogram (EOG) data using passive electrodes, with a ground electrode placed on the left cheek. Horizontal EOG was recorded from a bipolar pair of electrodes placed ~1 cm from the external canthus of each eye. Vertical EOG was recorded from a bipolar pair of electrodes placed above and below the right eye. Data were filtered online (low cut-off = .01 Hz, high cut-off = 80 Hz, slope from low- to high-cutoff = 12 dB/octave), and were digitized at 500 Hz using BrainVision Recorder (Brain Products, Munich, German) running on a PC. Impedance values were brought below 10 kΩ at the beginning of the session.

Eye tracking

We monitored gaze position using a desk-mounted EyeLink 1000 Plus infrared eye-tracking camera (SR Research, Ontario, Canada). Gaze position was sampled at 1000 Hz. According to the manufacturer, this system provides spatial resolution of .01° of visual angle, and average accuracy of .25-.50° of visual angle. We calibrated the eye tracker every 1-2 blocks of the task, and between trials during the blocks if necessary.

We drift-corrected the eye tracking data for each trial by subtracting the mean gaze position measured during a 200 ms window immediately before the onset of the memory array.

Artifact rejection

We segmented the EEG data into epochs time-locked to the onset of the memory array (200 ms before until 1000 ms after stimulus onset). We baseline-corrected the EEG data by subtracting mean voltage during the 200-ms window immediately prior to stimulus onset. Eye movements, blinks, blocking, drift, and muscle artifacts were first detected by applying automatic criteria. After automatic detection, we visually inspected the segmented EEG data for artifacts (amplifier saturation, excessive muscle noise, and skin potentials), and the eye tracking data for ocular artifacts (blinks, eye movements, and deviations in eye position from fixation), and discarded any epochs contaminated by artifacts. In all 3 experiments, all subjects included in the final sample had at least 200 trials of each set size condition (800 trials total).

Eye movements

For eye-tracking data, we rejected trials which contained eye movements beyond a certain threshold (threshold = 1° of visual angle). For some subjects, eye tracking data was not available (Exp. 1, n = 2; Exp. 2, n = 3). In these cases, EOG data was used. We rejected trials that contained horizontal or vertical EOG values beyond a threshold of 50 μV.

Blinks

In addition to the threshold detection, blinks were detected by flagging trials with flatline data (no position data are recorded when the eye is closed). Additionally, we visually inspected the eye tracking data for trial segments with missing data points.

Drift, muscle artifacts, and blocking

We checked for drift (e.g., skin potentials) with the pop_rejtrend function in ERPLAB. We excluded trials where a line fit to the EEG data has a slope greater than a certain threshold (slope = 10, minimal r^2 = .3). We checked for muscle artifacts with the pop_artmwppth function in ERPLAB. We excluded trials with peak to peak activity greater than 100 µV within a 200 ms window with 100 ms steps. We also excluded trials with any value beyond a threshold of 80 µV.

Multivariate Load (mvLoad) Procedure

Binned trial classification (within-subject & within-experiment)

The mvLoad analysis is within-subject classification of working memory load on baselined EEG. Although our approach allowed robust above-chance performance with single trials, we used randomly chosen groups of 20 trials within each set size to increase signal-to-noise ratio. We divided each trial into 50 ms windows with 25 ms steps and calculated the average voltage for each electrode in the window. Classification was performed using an ordinal logistic regression model. The classifier was trained to discriminate between load conditions 1, 2, 3, and 4 giving a chance level

classification of 25%. Classification was tested on a held out set of data using the StratifiedShuffleSplit function from Scikit-Learn. This cross-validation procedure splits the data in 80% training and 20% testing sets, while preserving the percentage of samples for each load condition. This split was repeated 1000 times and results for each subject and timepoint were averaged across these repetitions. Training data was standardized at each timepoint using the StandardScaler Scikit-Learn function, and test data was standardized using the mean and standard deviation of the training set.

Binned trial classification (within-subject & across-experiment)

Cross-training classification was used to test for generalization between the color (Exp 1) and orientation (Exp 2) conditions, and between the single-feature (Exps 1 and 2) and conjunction (Exp 3) conditions. These analyses followed the same procedures as the within-experiment classification except the testing was done on EEG data from a different experiment. For the single-feature generalizability analysis, the classifier was trained on data from Experiment 1 and tested on data from Experiment 2 and vice versa. For the single-feature to conjunction generalizability analysis, the classifiers were trained on a mixture of data from Experiment 1 and 2 and tested on data from Experiment 3. All of these analyses were done within-subject, using the subset of subjects who completed all of the experiments involved in the analysis (Exp. 1 & 2, n = 24; Exp. 1, 2, & 3, n = 20).

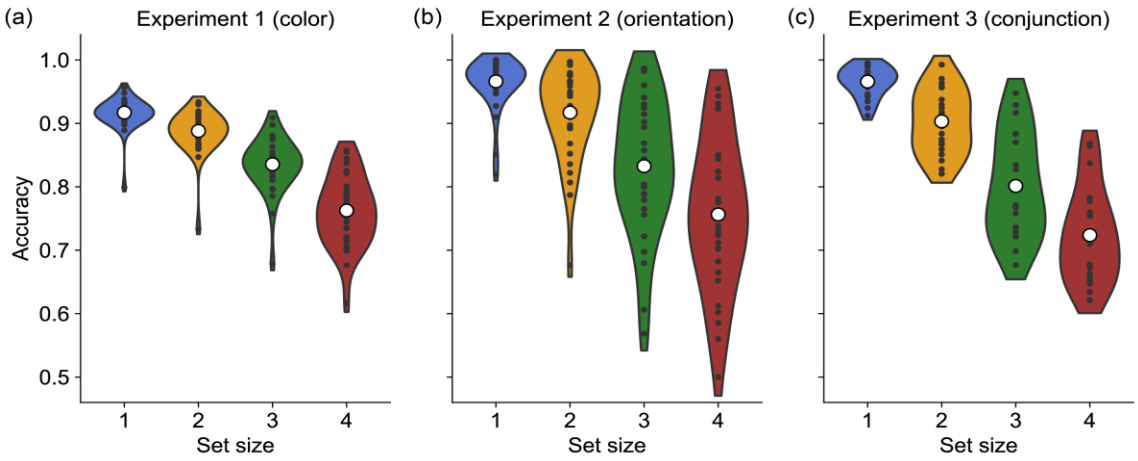

Figure 2 Change detection accuracy for each set size in each of the three experiments. See Supp. Fig. 2 to see Cowan's k for these data.

Significance Testing

In all classification analyses, we tested if classification accuracy was significantly above chance at each timepoint using a paired, one-tailed *t*-test. Classification accuracy was compared to empirical chance accuracy, defined by testing the trained model on randomly shuffled trial labels. Because we tested for significance at each timepoint (48 time bins between 0 and 1250 ms), we used the Benjamini-Hochberg procedure to control the false discovery rate (FDR) at .05.

Results

Behavioral

Across all experiments and conditions, subjects performed the change detection above chance (Figure 2, range of condition accuracies: .72-.97). In each experiment, a one-way ANOVA revealed a significant main effect of set size, such that accuracy declined as set size increased (Experiment 1, $F(3, 128) = 90.19$, $p < .001$; Experiment 2, $F(3, 120) = 32.93$, $p < .001$; Experiment 3, $F(3, 76) = 60.27$, $p < .001$). To examine whether behavioral performance varied across the 3 experiments, we carried out a within-subject

analysis using only the 20 observers who completed all 3 experiments. We combined data from Experiments 1 and 2 (single-feature items) to compare to Experiment 3 (conjunction items). In a two-way repeated measures ANOVA on accuracy, there was no significant main effect of feature, $F(1, 19) = 4.09$, $p = .057$, a significant main effect of set size, $F(3, 57) = 213.30$, $p < .001$, and a significant interaction of feature and set size, $F(3, 57) = 17.69$, $p < .001$. To characterize the significant interaction, we conducted 4 paired t-tests between the single feature and conjunction conditions at each set size (corrected p, FDR = .05 with Benjamini-Hochberg procedure). Set size 1 single-feature accuracy (M = .94, SD = .03) was significantly lower than conjunction (M = .97, SD = .02), $t(19) = -5.23$, $p < .001$, $d = .955$; set size 2 single-feature (M = .90, SD = .04) was not significantly different than conjunction (M = .90, SD = .05)), $t(19) = .001$, $p = .993$, $d = .050$; set size 3 single-feature accuracy (M = .84, SD = .06) was significantly higher than conjunction (M = .80, SD = .08)), $t(19) = 3.18$, $p = .007$, $d = .514$; and set size 4 single-feature accuracy (M = .76, SD = .08) was significantly higher than conjunction (M = .72, SD = .07)), $t(19) = 3.81$, $p = .002$, $d = .651$. Despite revealing reliably worse performance in the conjunction experiment, this still provides evidence for “object-based benefits” for storage in visual WM (Olson & Jiang, 2002). That is, a larger number of

feature values were stored in the conjunction condition than in the single-feature condition.

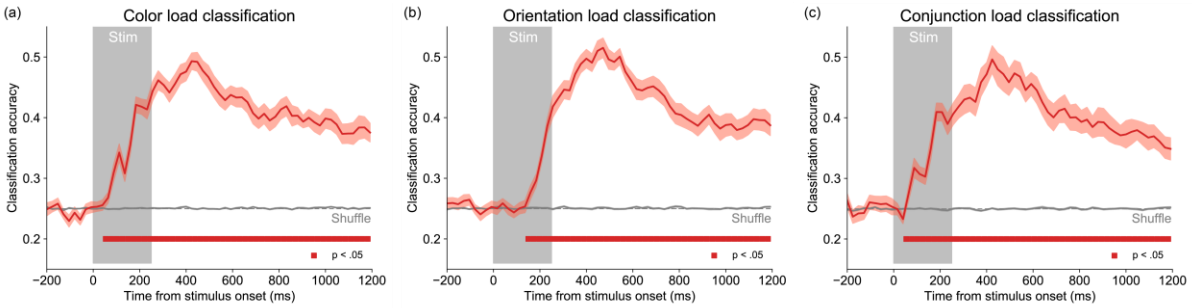

Figure 3 Classification accuracy over time for each experiment. Classification accuracy is indicated with a red line. Shaded error bars indicate ± SEM. Red squares indicate timepoints with classification significantly above chance (corrected p < .05, FDR = .05 with Benjamini-Hochberg procedure). Grey line indicates chance classification accuracy. Grey rectangle indicates time period where memory array is displayed.

Precise classification of load while controlling for stimulus energy

The first key result was that the mvLoad analysis precisely classified working memory load, despite the use of stimulus displays that controlled for stimulus energy across all load conditions. For each experiment (Exp. 1, n = 33; Exp. 2, n = 31; Exp. 3, n = 20), we used an ordinal logistic regression classifier on raw EEG amplitudes from binned trials within-

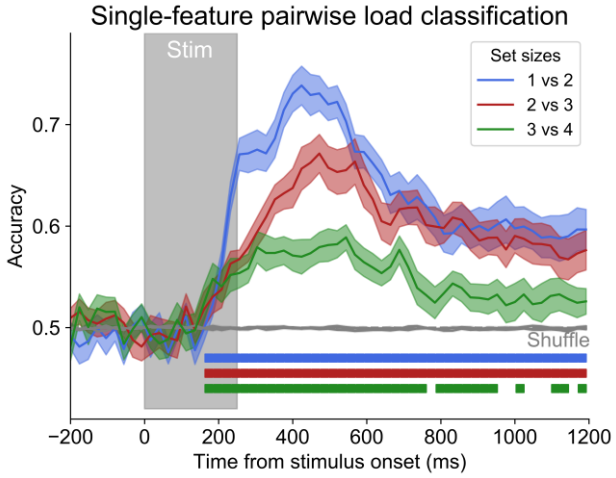

Figure 4 Classification accuracy of single-feature load (experiments 1 and 2 mixed together) for set sizes 1 vs. 2, 2 vs. 3, and 3 vs. 4. Colored lines indicate classification accuracy. Shaded error bars indicate ± *SEM.* Color-matched squares indicate timepoints with classification significantly above chance (corrected p < .05, *FDR* = .05 with Benjamini-Hochberg procedure). Grey line indicates chance classification accuracy. Grey rectangle indicates time period where memory array is displayed.

subject (20 trials per bin) at each time bin (50 ms window). We could classify working memory load (set size 1 vs. 2 vs. 3 vs. 4) during the stimulus presentation and throughout the delay period (Figure 3a, 3b, 3c; red squares indicate corrected $p < .05$, FDR-controlled at .05 with Benjamini-Hochberg procedure with 48 time-bins tested). Above-chance classification was observed starting in early time bins in each experiment (Exp. 1, 64-88 ms timebin; Exp. 2, 160-208 ms timebin; Exp. 3, 64-88 ms timebin). Classification was sustained throughout the entire delay period for all three experiments. Mean classification accuracy (with chance at .25) during the delay period for Experiment 1 was .42 (SD = .03); Experiment 2, .43 (SD = .04); Experiment 3, .41 (SD = .04). We also confirmed that the classifier was sensitive to single item increments in the number of stored items. Figure 4 shows classification accuracy for set sizes 1 vs. 2, 2 vs. 3, and 3 vs. 4. For 1 vs. 2 and 2 vs. 3, accuracy was sustained above chance throughout the entire delay period (corrected $p < .05$, *FDR* = .05 with Benjamini-Hochberg procedure). Using behavioral (set size 4) and EEG data from each unique subject across the three experiments (N = 40), we replicated the findings (from Adam et al., 2020; Feldmann-Wüstefeld, 2021) that classification accuracy was positively correlated with individual differences in working memory capacity ($r^2 = .24$, $p = .001$) (Figure 5). Further analysis shows that this relationship is

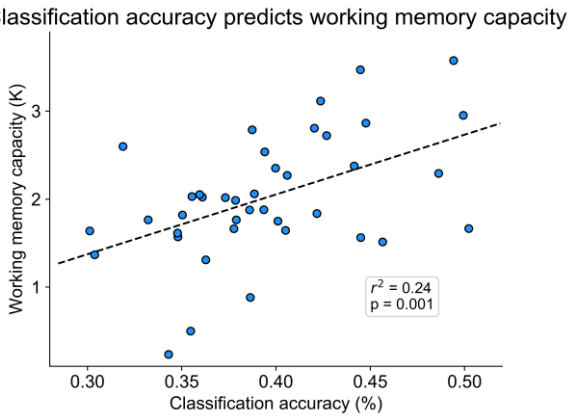

Figure 5 Individual differences in classification accuracy are positively correlated with working memory capacity. Classification accuracy is average delay period accuracy from all unique subjects across all three experiments; all data was used from each unique subject. Working memory capacity was measured using set size 4 trials.

consistent across nearly all timepoints in the delay period. This correlation may be caused by the greater reliability with which higher capacity individuals achieve the storage of all relevant items (e.g., Adam et al., 2015), which would in turn yield more discriminable patterns of activity for each set size. This finding reinforces the earlier evidence that the mvLoad analysis taps into a neural operation that is relevant for understanding capacity limits in visual working memory.

A load signature that generalizes across distinct feature values

The second key analysis examined whether the load signatures revealed by mvLoad *generalized* across distinct feature values (i.e., color and orientation). Using data from subjects who had participated in both Experiments 1 and 2 (n = 24) we trained the classifier using the color trials from Experiment 1 and tested it on the orientation trials from Experiment 2. We also trained on orientation trials and tested on color trials. In both directions of

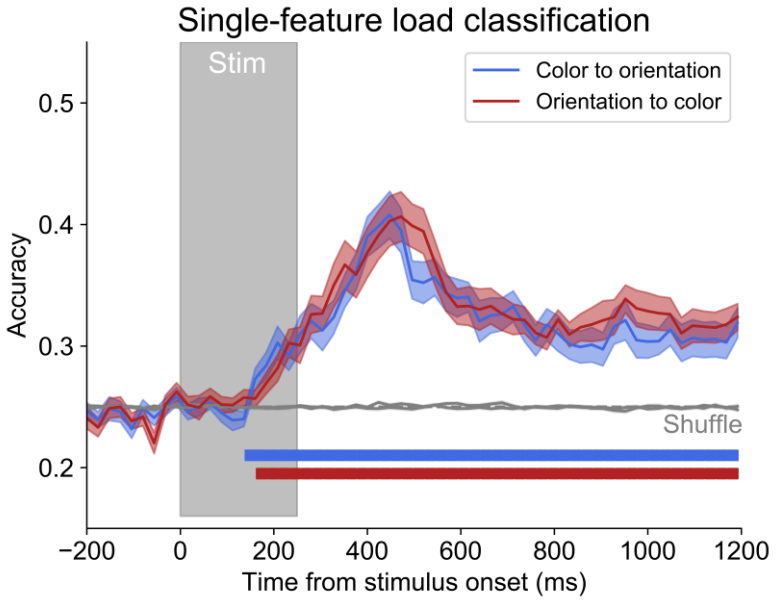

Figure 6 Accuracy for single-feature load classification. Blue line is classification accuracy when trained on data from experiment 1 (color) and tested on experiment 2 (orientation). Red line is accuracy when trained on experiment 2 and tested on experiment 1. Color-matched squares indicate timepoints with classification significantly above chance (corrected $p < .05$, *FDR* = .05 with Benjamini-Hochberg procedure). Grey line indicates chance classification accuracy. Grey rectangle indicates time period where memory array is displayed.

training and testing, robust classification was sustained throughout the entire delay period (Figure 6). Mean classification accuracy during the delay period for color to orientation was .33 (SD = .03); orientation to color, .34 (SD = .03). Thus, the same multivariate pattern classified load precisely for memoranda with distinct relevant features, revealing a load-sensitive signal that is separable from the specific content stored in working memory. These decoding accuracies are lower than we saw with within-experiment analyses. While this could reflect non-generalizable aspects of the load signal, it could also reflect methodological noise across sessions, such as small differences in electrode placement or impedence. Thus, even if precisely the same load pattern were present in each EEG session, some drop in decoding accuracy would be expected for across-session relative to within-session training.

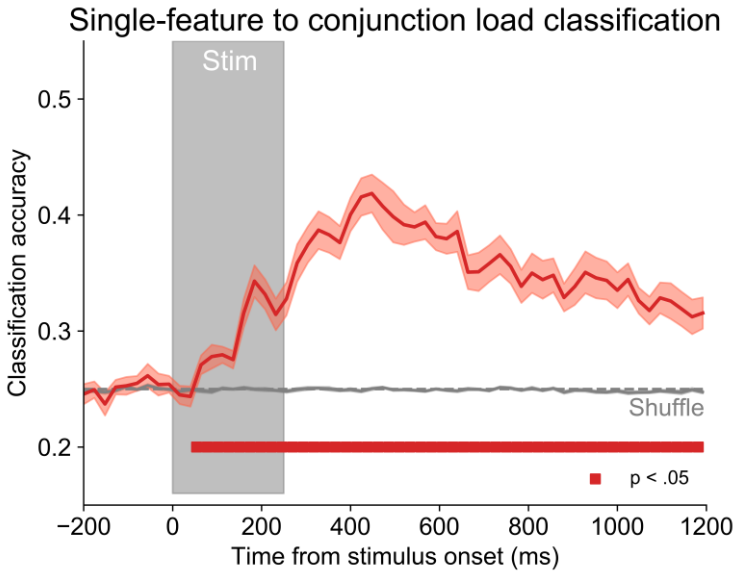

Figure 7 Accuracy for classification trained on data from experiments 1 and 2 (single-feature items, color *or* orientation) and tested on data from experiment 3 (conjunction items, color *and* orientation). Red squares indicate timepoints with classification significantly above chance (corrected $p < .05$, *FDR* = .05 with Benjamini-Hochberg procedure). Grey line indicates chance classification accuracy. Grey rectangle indicates time period where memory array is displayed.

A signature of load that is independent of total amount of information stored

The third key analysis examined whether the load-sensitive activity revealed by the mvLoad analysis was independent of the total amount of feature information maintained about each item stored in working memory. To this end, we trained the classifier on the combined data from Experiments 1 and 2, in which each item contained one relevant feature to be stored (i.e., either color or orientation) and we tested this model using data from Experiment 3 in which the number of relevant features per item was doubled (i.e., both color and orientation) (Figure 7). This analysis included a group of 20 subjects who had participated in all three experiments. Classification accuracy was robustly above chance throughout the entire delay period with a mean delay-period accuracy of .36 (SD = .03). Again, the across-experiment decoding accuracy was lower than in the within-experiment analyses. Nevertheless, the same signature of load identified with single feature stimuli was observed with conjunction stimuli that contained twice as many relevant features per item, in line with a load-sensitive cognitive operation that is separable from the maintenance of specific features.

While robust cross-training between single feature and conjunction conditions suggests that they evoked a common load signature, further analyses provided more incisive evidence for the content-independent character of this load-sensitive neural activity. First, recall from Experiments 1 and 2 that the mvLoad analysis robustly detected the difference between one and two single-feature items (Figure 3), and between two and three single-feature items, showing that the analysis is sensitive to the addition of a single item with one relevant feature. Thus, if load decoding with single-feature items was based on the number of color or orientation values stored, then 1 conjunction item

should be classified as the same load as 2 single-feature items. Alternatively, if load decoding was based on the number of feature-independent pointers stored, then 1 conjunction item should be classified as the same load as 1 single-feature item. To test this prediction, we trained the mvLoad classifier with single-feature stimuli, and

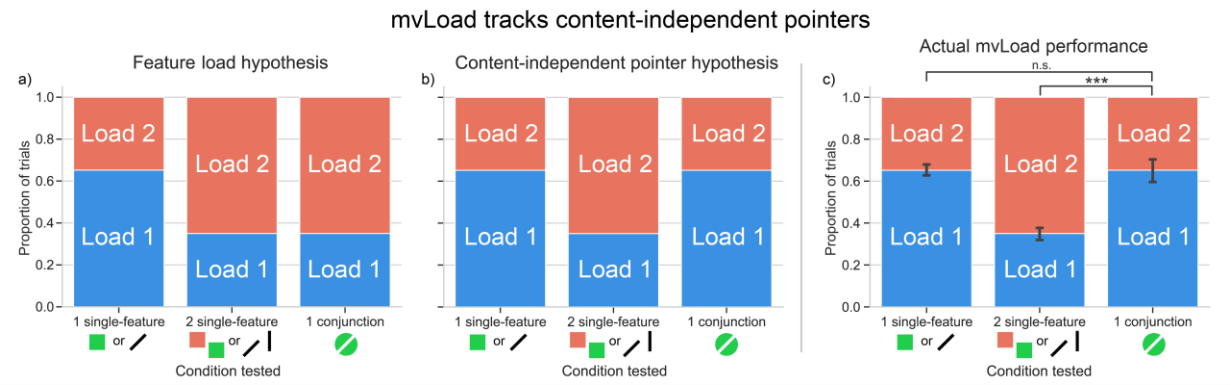

Figure 8 mvLoad classifier output (trained on single-feature items) compared to the predicted output of the feature load and pointer hypotheses. Within each graph, each bar represents 100 percent of trials classified as load 1 (blue) or 2 (red) in each condition. (a) The feature load hypothesis predicts that load 1 conjunction items will be classified the same way as load 2 single-feature items. (b) The pointer hypothesis predicts that load 1 conjunction items will be classified the same way as load 1 single-feature items. (c) The actual mvLoad classifier performance was in-line with the pointer hypothesis.

examined performance across 3 key conditions: (1) set size 1 single-feature (2) set size 2 single-feature (3) set size 1 conjunction. The divergent predictions of the feature-load and pointer explanations are illustrated in Figure 8a and 8b, along with the observed data in 8c.

Visual inspection reveals that our findings fall directly in line with the pointer hypothesis, such that a one conjunction item was equivalent to one single-feature item. We tested the reliability of this pattern with two planned comparisons. First, we found a reliable difference between the predicted load for a single conjunction item and two single-feature items. A Bayesian paired t-test revealed strong evidence for a difference

between these conditions ($t(19) = -9.01$, $p < .001$, $d = 3.09$, $BF_{10} > 100$), showing that a single conjunction item had a higher probability to be classified as load 1 (M = .65, SD = .12) than set size 2 single-feature items (M = .35, SD = .07). Second, we examined the prediction that one conjunction item should predicted as the same load as one single-feature item (M = .65, SD = .06), using a Bayesian paired t-test. This revealed substantial evidence for the null hypothesis, suggesting that both had the same probability of being classified as load 1 ($t(19) = 0.00$, $p = .999$, $d = 0.00$, $BF_{10} = .232$). An analogous analysis of set size 2 and 4 trials revealed precisely the same empirical pattern, showing that 2 conjunction items (M = .68 classified as load 2, SD = .15) were predicted as a lower load than 4 single-feature items (M = .34, SD = .08), ($t(19) = -8.06$, $p < .001$, $d = 2.76$, $BF_{10} > 100$), and that 2 conjunction items were predicted as the same load as 2 single-feature items (M = .66, SD = .08), ($t(19) = -.652$, $p = .522$, $d = .185$, $BF_{10} = .281$). Thus, our findings strongly suggest that there is a common load signature for single-feature and conjunction stimuli that is determined by the number of individuated items stored, rather than by the number of feature values stored.

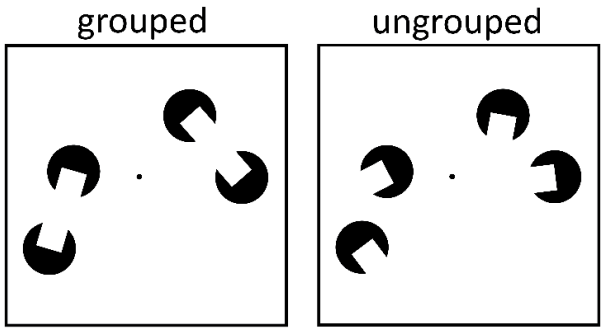

Figure 9 Examples of a set size 4 grouped and ungrouped memory arrays. In the grouped condition, collinearity between the notches yields the percept of a single oriented rectangle for each grouped pair.

Ruling out the size of the attended region as the driver of load-sensitive neural activity

Although the results of the mvLoad analysis point to load-sensitive neural activity that is separable from the quantity and type of content stored about each item, we noted that the *spatial extent* of the attended region in the display was confounded with the number of stored items. Thus, we examined whether the classifier was indexing the area of the attended regions on the screen, rather than the number of individuated items, per se. To this end, we re-analyzed data from an EEG study of perceptual grouping by Diaz et al. (2021) in which subjects stored the orientation of two or four notched discs in visual working memory. In the *grouped* condition, the discs were arranged so that collinearity between the notches in pairs of discs elicited the percept of a single illusory rectangle (Figure 9). Thus, in the set size 4 grouped condition, perceptual grouping encouraged the perception of *two* individuated orientation values, whereas in the set size 4 ungrouped condition observers perceived *four* individuated orientation values. Critically, the number of relevant elements and their spatial extent were matched between the grouped and ungrouped displays. Diaz et al. (2021) reinforced this point by showing that the power of alpha oscillations in occipitoparietal electrodes, a neural signal that has been shown to track the number of attended locations (Fukuda & Woodman, 2015), tracked the number of elements on the screen, but was unaffected by the grouping manipulation. Thus, the key question for the present study was whether the mvLoad classifier would register the difference between the grouped and ungrouped displays. If load classification is based on the spatial extent of the attended locations, then it should return the same load value for the grouped and ungrouped conditions, in line with the posterior alpha power signal examined by Diaz et al. (2021). By contrast, if load classification is based on the number of individuated items stored, then a lower load

should be detected in the grouped relative to the ungrouped condition. Figure 10 illustrates the output over time of a classifier that was trained exclusively on ungrouped displays (set size 2 or 4) and then tested on both the ungrouped and grouped displays. The output here is from the classifier's decision_function method, which returns the confidence score of the sample. This score is proportional to the signed distance of that sample to the hyperplane. In Figure 10, stronger evidence for set size 4 is plotted in the positive direction, whereas stronger evidence for set size 2 is plotted in the negative direction. When trained and tested on the ungrouped trials, the classifier exhibited sustained above-chance performance throughout the delay period (i.e., sustained positive values for set size 4 and sustained negative values for set size 2). However, when the same classifier was tested with set size 4 *grouped* trials, classification evolved over time. Set size 4 grouped trials were initially classified the same as set size 4 ungrouped. However, by the 512-536 ms time bin, set size 4 grouped diverged away from set size 4 ungrouped, and was reliably

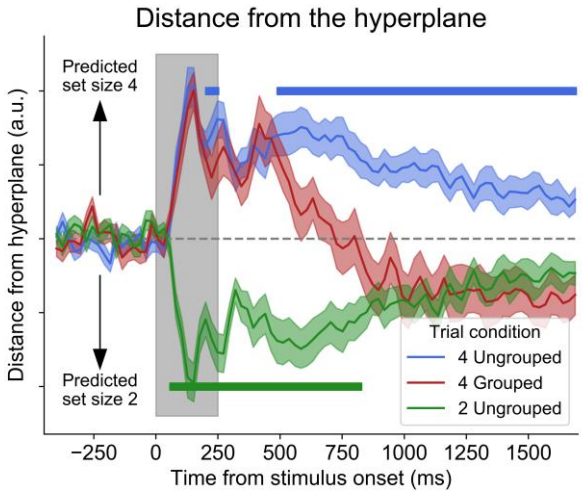

Figure 10 Distance from the classification hyperplane for set size 2 ungrouped, set size 4 grouped, and set size 4 ungrouped trials across time. Classification is trained on set size 2 and 4 ungrouped trials and tested on all three conditions. Hyperplane is indicated with the dashed grey line. Trials above the hyperplane are classified as set size 4 while trials below it are classified as set size 2. Distance from the hyperplane for each trial condition at each timepoint. Blue squares indicate timepoints where 4 ungrouped is significantly greater than 4 grouped (corrected $p < .05$, *FDR* = .05 with Benjamini-Hochberg procedure). Green squares indicate where 2 ungrouped is significantly less than 4 grouped (corrected $p < .05$, *FDR* = .05 with Benjamini-Hochberg procedure).

closer to the hyperplane. Set size 4 grouped was also reliably different from set size 2 ungrouped at the start of the trial. However, by time bin 848-872 ms, set size 4 grouped had crossed the hyperplane and was no longer reliably different from set size 2 ungrouped. Thus, while perceptual grouping did not affect the spatial extent of the attended region (Diaz et al., 2021), the mvLoad classifier indexed a lower number of stored items in the grouped condition, showing that the classifier indexes the number of individuated items stored in memory, not the spatial extent of covert attention.

Discussion

Given that working memory serves as a cornerstone for intelligent behaviors, there is strong motivation to build a taxonomy of the neural operations that support online memory storage. The dominant strain of this work has focused on stimulus-specific neural activity that represents the stored content (e.g., D'Esposito & Postle, 2015; Funahashi et al., 1989; Fuster & Jervey, 1981; Goldman-Rakic, 1995; Rademaker et al., 2019; Serences et al., 2009), and great progress has been made in understanding the format and anatomical locus of this class of neural activity. By contrast, we highlight evidence for a qualitatively different neural operation that is integral to WM function, but separable from the maintenance of stored content. Specifically, we refer to a spatiotemporal pointer operation that supports the segmentation of visual scenes into individuated representations that can be tracked through time and space (Khaneman et al, 1992; Pylyshyn, 2009). Using a multivariate analytic approach (Adam et al., 2020), we show that the scalp topography of EEG voltage precisely tracks the number of individuated representations stored in visual WM, while generalizing across variations in

both the type and number of relevant features per item. Thus, although this neural operation is hypothesized to track the spatiotemporal coordinates of stored objects, it operates in a fashion that is insensitive to the *contents* of the tracked memory representations. Moreover, the fidelity of this load-sensitive neural activity is a predictor of individual differences in WM capacity, emphasizing its importance for understanding *why* WM capacity is limited.

The present findings provide a critical complement to past work that has sought to determine the computational role of load-sensitive neural activity. For instance, multiple studies have reported EEG and BOLD activity patterns that rise with each additional item stored, and reach an apparent plateau at set sizes that exceed behavioral estimates of capacity in visual WM (e.g., Todd & Marois, 2004; Vogel & Machizawa, 2004; Xu & Chun, 2006). But while this empirical pattern is consistent with a neural operation that tracks number *per se*, it can also be modeled using a biophysically plausible “saturation model” wherein stimulus-specific neural activity follows an exponential function (Bays, 2018). There have also been reports of neural activity that rises with the number of items, but that isn’t affected by the complexity of the memoranda (Xu and Chun, 2005; Woodman and Vogel, 2008). This empirical pattern suggests a neural operation that indexes the number of individuated representations stored in working memory rather than the total amount of visual information. That said, these conclusions are based on an intriguing null result: the absence of a difference in mean activity levels across distinct types of stimuli. By contrast, our findings provide positive evidence for a common neural index of the number of stored items when the

type and number of visual features per item is varied: a multivariate signature of load that robustly generalizes across three distinct types of memoranda, demonstrating a content-independent aspect of storage-related neural activity. Moreover, our findings were supported by 84 separate EEG sessions across 40 unique observers that yielded above-chance decoding in every session for every observer tested. Thus, our findings provide compelling positive evidence for an item-based, content-independent aspect of storage in visual working memory.

Our working hypothesis is that this load-sensitive neural activity reflects the deployment of spatiotemporal “pointers” or indexes that support object individuation -- the segmentation of objects from the background and from other objects -- and the continuous tracking of items through time and space (e.g., Pylyshyn, 2009; Khaneman et al., 1992; Xu and Chun, 2009). To study this cognitive process, Pylyshyn & Storm (1988) introduced “multiple object tracking” (MOT), a task that requires the observer to keep track of varying numbers of targets that move randomly amongst a group of identical distractors. Their behavioral data indicated a relatively sharp capacity limit that they attributed to a limit on the number of pointers that could be concurrently deployed. Interestingly, Drew and Vogel (2009) used a lateralized version of the MOT task to show that CDA activity rises with the number of targets that are tracked, predicts individual differences in tracking ability, and reaches an apparent plateau after three targets are selected. Thus, it may be that both the CDA and mvLoad classifiers are picking up on a content-independent indexing operation that is required during tracking and visual working memory tasks (Balaban et al., 2019; Hakim et al., 2019; Tsubomi et al., 2013).

In combination with stimulus-selective neural activity that supports the maintenance of precise memories (e.g., D'Esposito & Postle, 2015), evidence for a content-independent pointer operation falls in line with various proposals for a separation between the precise maintenance of content and the number of representations maintained in working memory. For example, if WM storage is limited by the deployment of content-independent pointers, this could explain why the maximum number of items an individual can store is uncorrelated with the precision of those representations (Awh et al., 2007), and why number exhibits a strong correlation with fluid intelligence while precision does not (Fukuda et al., 2010). Likewise, this separation may explain why different regions of visual cortex appear to track the number and complexity of the memoranda stored in WM (e.g., Xu and Chun, 2005). In addition, if storage in visual WM is contingent on the assignment of a pointer, this could explain why many studies have documented an "object-based benefit" in which a larger number of features can be maintained within multi-feature compared to single-feature objects (Luck & Vogel, 1997; Olson & Jiang, 2002). Specifically, if each individuated object stored requires one of a limited number of pointers, then single-feature items would be the least efficient way to store the largest number of features.

In conclusion, multivariate analysis of the topography of EEG voltage reveals a load-sensitive neural operation that tracks the number of individuated items stored in working memory, while generalizing across variations in the type and number of visual features. This empirical pattern provides critical new evidence for a distinction between the maintenance of visual features and the discrete indexing of the items that contain those

features. These findings help to clarify the taxonomy of neural operations that support storage in this online mental workspace.

Chapter 2: Stronger evidence for feature-independent pointers in visual working memory

Introduction

Visual working memory (WM) research has focused a great deal on stimulus-specific neural activity that supports the maintenance of information. There is significant evidence that WM elicits sustained neural activity in populations that are selective for the particular features values held in mind (D'Esposito & Postle, 2015; Funahashi et al., 1989; Fuster & Jervey, 1981; Goldman-Rakic, 1995; Harrison & Tong, 2009; Rademaker et al., 2019; Serences et al., 2009). These studies can provide insight into the neural activity that underlies the maintenance of specific information in mind.

Another area of WM research is neural measures of load, or, the amount of information held in mind (Adam et al., 2020; Todd & Marois, 2004; Vogel & Machizawa, 2004; Xu & Chun, 2009). For example, Vogel and Machizawa 2004 discovered a contralateral negativity in EEG for items stored in WM. This contralateral delay activity (CDA) has since been shown to track the number of objects in WM (Ikkai et al., 2010) and is highly predictive of individual WM capacity.

A long-standing discussion of WM storage relates to the process of binding or individuating objects. Kahneman et al 1992 proposed the object file as a method of binding an object as a temporary episodic representation to support tracking of the object through time and space. (Kahneman et al., 1992) Pylyshyn 1989 similarly presents the

concept of the “finger of instantiation” (FINST), which are indexes of individuated objects separate from their type or location. (Z. Pylyshyn, 1989) Both theories describe a type of spatio-temporal pointer that supports the tracking and storage of objects while the actual content is maintained by a separate mechanism.

There have been several models of WM that propose separate neural processes for maintenance of features and individuation (Balaban et al., 2019; Bouchacourt & Buschman, 2019; Oberauer, 2019; Swan & Wyble, 2014; Xu & Chun, 2009). For example, Xu and Chun (2009) argued for two separate neural pathways for object identification and individuation. Swan and Wyble (2014) proposed a “neural binding pool” which binds together multiple features of an object serving the representation as an individuated token.

Adam et al 2020 used multivariate classification of EEG activity to decode the number of objects stored in WM. (Adam et al., 2020) This multivariate load detection (mvLoad) is an incredibly sensitive measure, even working on individual EEG trials. It also generalized across individuals and unique task designs (whole field versus lateralized memory displays). Chapter 1 of this dissertation presented evidence that mvLoad generalizes across items with distinct feature values and from single-feature to conjunction items (Thyer et al., 2022). It was also not affected by the spatial extent of the objects in memory. This strongly suggests that this mvLoad procedure is tracking content-independent spatio-temporal pointers.

However, a key limitation of those experiments is that color and orientation change detection tasks were collected in different sessions. A drop in mvLoad generalization could be solely due to session noise causing the feature-independence of the model to be underestimated. Another limitation that could cause an overestimation of feature-independence is that neural populations which represent colors and orientations may be so interdigitated that they don't produce distinct EEG signals. This would partially undermine the cross-training results. By using more separable populations, such as color and motion coherence, (Felleman & Van Essen, 1991; Vaina, 1994; Zeki, 1978) we can more accurately test the feature generalizability of mvLoad.

If mvLoad generalization between color and orientation data that was collected within one session still shows a drop in accuracy compared to within-feature classification, then that is evidence for the presence of some feature-specific activity. However, if load classification generalizes perfectly, then we have strong evidence for a feature-independent load signature. With color and motion objects, we have an even stronger test of this feature-generalizability. If there is no cross-training between color and motion objects, then that suggests that there is indeed a strong limitation to the feature-generalizability of mvLoad. But, if there is above-chance classification, even in objects with features coded in very disparate neural populations, then we have strong evidence for feature-independence which is in line with the hypothesis that mvLoad is tracking feature-independent pointers in WM.

Methods

Subjects

Experiments included 35 volunteers (Experiment 1, n = 13; Experiment 2, n = 15) participating for monetary compensation ($15 per hr). Subjects were between the ages of 18 and 35 years old, reported normal or correct-to-normal visual acuity, and provided informed consent according to procedures approved by The University of Chicago Institutional Review Board. Subjects were recruited via online advertisements and fliers posted on the university campus.

Experiment 1. Our target sample in Experiment 1 was 12 subjects. 17 volunteers participated in Experiment 1 (8 females; mean age = 24.9 years, SD = 3.8). 4 subjects were excluded from the final sample for the following reasons: The session was ended early due to eye movements (n = 2); the subject's data was corrupted or otherwise unusable (n = 2). The final sample size was 13 (6 female; mean age = 25.0 years, SD = 4.1).

Experiment 2. Our target sample in Experiment 1 was 15 subjects. 18 volunteers participated in Experiment 2 (9 females; mean age = 26.6 years, SD = 4.0). 3 subjects were excluded from the final sample for the following reasons: The session was ended early due to eye movements (n = 2); the subject didn't have enough data after artifact rejection (n = 1). The final sample size was 15 (9 female; mean age = 26.7 years, SD = 2.2).

Apparatus

We tested the subjects in a dimly lit, electrically shielded chamber. Stimuli were generated using PsychoPy (Peirce et al., 2019). Subjects viewed the stimuli on a gamma-corrected 24-in. LCD monitor (refresh rate = 120 Hz, resolution = 1,080 × 1,920 pixels) with their chins on a padded chin rest at a viewing distance of 75 cm.

Luminance-balanced displays

Experiment 1. Stimuli were presented against a mid-gray background (~61 cd/m^2). Memory arrays included one or three to-be remembered items. Ignored placeholder items also appeared in the memory array, so each array had a total of four items. The placeholder items were shown in a shade of gray (red, green, blue [RGB] value = 149, 150,149) that matched the average luminance of all possible colors in the color set. Placeholders were circles (radius = 1.13°) to match the area of memory items.

Experiment 2. Stimuli were presented against a mid-gray background (~61 cd/m^2). Memory arrays included one or two to-be remembered items. Ignored placeholder items also appeared in the memory array, so each array had a total of three items. The placeholder items were shown in a shade of gray (red, green, blue [RGB] value = 166, 166, 166) that matched the average luminance of all possible colors in the color set. Ignored placeholders were motion coherence patches of the same speed, size, and number of dots as the memory items, but they always had 0 motion coherence.

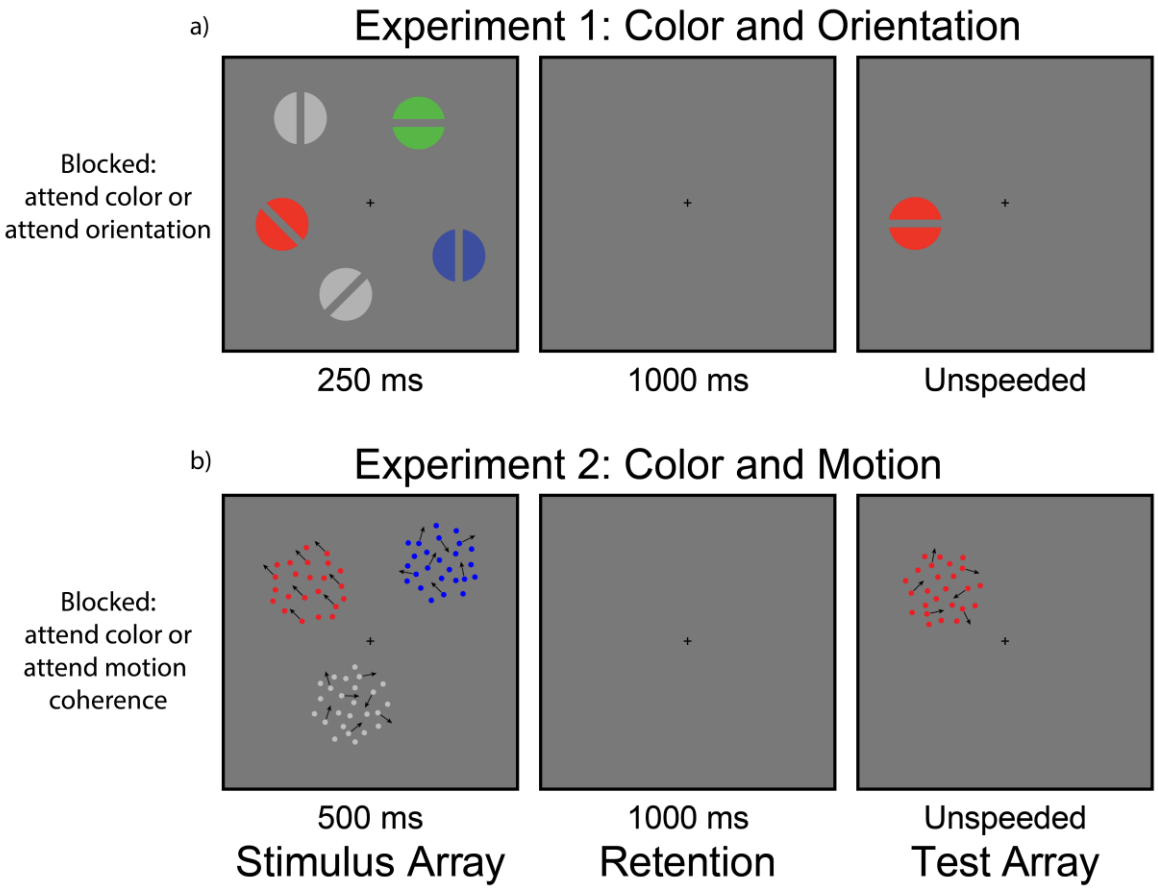

Figure 11 Task schematics for an example Set Size 3 trial in the whole-field change-detection task used on both experiments. In Experiment 1 (a), subjects were cued at the beginning of each block to either remember the color or orientation of the memory items while ignoring the gray distractor items. In Experiment 2 (b), subjects were cued at the beginning of each block to either remember the color or the motion coherence of the memory items while ignoring the gray distractor items. During change trials, the attended feature changes 50% of the time.

Task procedures

Experiment 1. The experiment used a whole-field change detection task (Figure 11a). On each trial, a memory array appeared containing four total items. There were one or four colored items to be remembered, and the remainder of the items were gray placeholder items to balance area and luminance across set-size conditions (see the *Luminance-balanced displays* section for more detail).

Memory items were colored circles (radius = 1.3°) with oriented bars cut out of the middle (height = 2.6°, width = 0.5°). The possible orientations were 0°, 90°, 180°, and 270°, and they were sampled without replacement for each trial. The possible colors were randomly sampled without replacement from a set of four colors (RGB values: red = 255, 0, 0; green = 0, 255, 0; blue = 0, 0, 255; yellow = 255, 255, 0). Items were positioned with a maximum of one item per quadrant. In each block, either the color or orientation was indicated as the target feature for that block. Two memory items never appeared in one quadrant together, and all items were placed at least 4 apart. Subjects viewed a memory array (250 ms), remembered items across a delay (1000 ms), were probed on one item, and reported whether the probed item was the same as or different from the remembered item (unspeeded). Alternating each block, participants were instructed to attend to either the color or the orientation of the memory items. Only the attended feature dimension could change.

Experiment 2. The experiment used a whole-field color report task (Figure 11b). On each trial, a memory array appeared containing 3 total items. Items were colored moving dot patches (radius = 3º). The possible colors were randomly sampled without replacement from the same set of seven colors as Experiment 1 (RGB values: red = 255, 0, 0; green = 0, 255, 0; blue = 0, 0, 255; yellow = 255, 255, 0; purple = 255, 0, 255; teal = 0, 255, 255; orange = 255, 128, 0). The moving dots within each item were either moving coherently (all in one direction) or incoherently (moving in random directions). Items were positioned with a maximum of one item per quadrant. In each block, either the color or the motion coherence was indicated as the feature to be remembered. Subjects viewed

a memory array (500 ms), remembered items across a delay (1000 ms), were probed on one item, and reported whether the probed item was the same as or different from the remembered item (unspeeded). Alternating each block, participants were instructed to attend to either the color or the motion coherence of the memory items. Only the attended feature dimension could change.

EEG acquisition

We recorded EEG activity from 30 active Ag/AgCl electrodes mounted in an elastic cap (Brain Products actiCHamp, Munich, Germany). We recorded from international 10-20 sites Fp1, Fp2, F7, F3, Fz, F4, F8, FT9, FC5, FC1, FC2, FC6, FT10, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, and O2. Two additional electrodes were affixed with stickers to the left and right mastoids, and a ground electrode was placed in the elastic cap at position Fpz. All sites were recorded with a right-mastoid reference and were rereferenced off-line to the algebraic average of the left and right mastoids. Data were filtered on-line (low cut-off = 0.01 Hz, high cut-off = 80 Hz, slope from low to high cut-off = 12 dB/octave) and were digitized at 500 Hz using BrainVision Recorder (Brain Products, Munich, Germany) running on a PC. Impedance values were brought below 10 kΩ at the beginning of the session.

Eye tracking

We monitored gaze position using a desk-mounted EyeLink 1000 Plus infrared eye-tracking camera (SR Research, Ontario, Canada). Gaze position was sampled at 1000 Hz. According to the manufacturer, this system provides spatial resolution of .01° of visual

angle and average accuracy of 0.25 to 0.50° of visual angle. We calibrated the eye tracker every one to two blocks of the task and between trials during the blocks if necessary. We drift-corrected every 5 trials. Additionally, we drift-corrected the eye-tracking data for each trial by subtracting the mean gaze position measured during a 200-ms window immediately before the onset of the memory array.

Artifact rejection

We segmented the EEG data into epochs time-locked to the onset of the memory array (200 ms before until 1,000 ms after stimulus onset). We baseline-corrected the EEG data by subtracting mean voltage during the 200-ms window immediately prior to stimulus onset. Eye movements, blinks, blocking, drift, and muscle artifacts were detected by applying automatic criteria and we discarded any epochs contaminated by artifacts. All subjects included in the final sample had at least 140 trials of each condition.

Eye movements and blinks

We employed real-time eye movement detection. If participants moved their eyes more than 1.25° from fixation, the trial was interrupted and their eye position was shown to them for feedback purposes. Interrupted trials were made up at the end of the block. During preprocessing, we rejected trials that contained eye movements beyond 1° of visual angle using the pop_artextval function in ERPLAB (Lopez-Calderon & Luck, 2014).

Drift and muscle artifacts

We checked for drift (e.g., skin potentials) with the pop_rejtrend function in ERPLAB.). We checked for muscle artifacts with the pop_artextval function in ERPLAB. We excluded trials where EEG activity was greater than 80 µV or less than -80 µV. We excluded trials with peak-to-peak activity greater than 100 µV within a 200-ms window with 100- ms steps. We also excluded trials with any value beyond a threshold of 80 µV.

mvLoad procedure

Load classification (within-attended feature)

The mvLoad analysis is within-subjects classification of WM load on baselined EEG. Although our approach allowed robust above-chance performance with single trials, we used randomly chosen groups of 20 trials within each set size and attended feature to increase signal-to-noise ratio. We divided each trial into 50-ms windows with 25-ms steps and calculated the average voltage for each electrode in the window. Classification was performed using the Scikit-Learn Logistic Regression model. The classifier was trained to discriminate between load conditions x and y within each attended feature, giving a chance-level classification of 50%. Classification was tested on a held-out set of data using the train_test_split function from Scikit-Learn (Pedregosa et al., 2011) stratified on the trial conditions. This cross-validation procedure splits the data in 80% training and 20% testing sets while preserving the percentage of samples for each load condition. This split was repeated 1,000 times, and results for each subject and time point were averaged across these repetitions. Training data were standardized at each time point using the

StandardScaler Scikit-Learn function, and test data were standardized using the mean and standard deviation of the training set.

Load classification (across-attended feature)

Cross-training classification was used to test for generalization between the color and orientation conditions. These analyses followed the same procedures as the within-attended feature classification except that the testing was done on EEG data from the other attended feature condition. The classifier was trained on data from the attend color condition and tested on data from the attend orientation condition and vice versa.

Attended feature classification

Attended feature was also classified using a similar procedure as mvLoad. The standard LogisticRegression from Scikit-Learn model was used instead of an ordinal logistic regression model. Also, the classifier is trained to discriminate between attended color and attended orientation conditions. Load conditions 1 and 3 are collapsed across within each attend feature group.

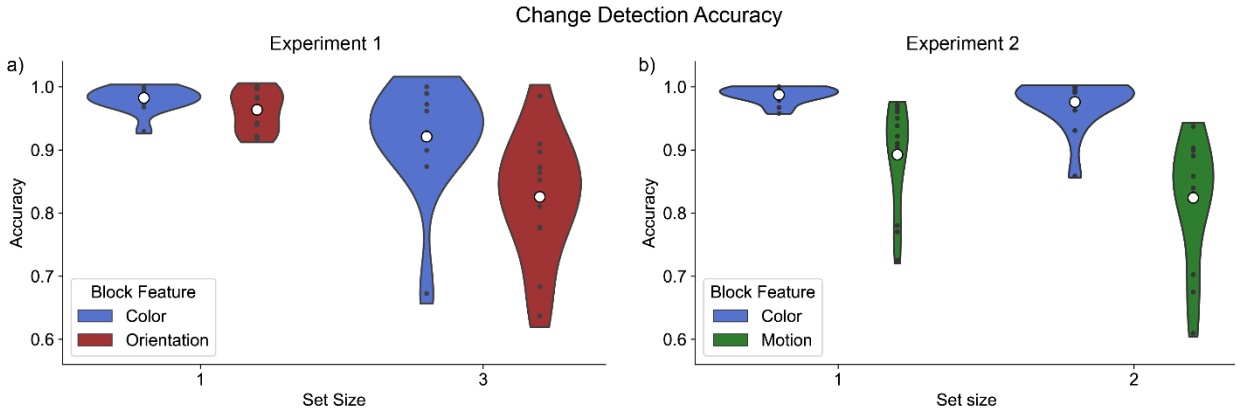

Figure 12 Change detection accuracy for each set size in both experiments. Black dots indicate individual data, white dots indicate means, and shaded regions indicate the density of the data.

Coherence level classification

Coherence level was also classified using a similar procedure as mvLoad. The LogisticRegression model from Scikit-Learn was trained to discriminate between coherent and incoherent memory objects in the set size 1 condition. We also compare coherence classification in the Attend Motion Coherence versus the Attend Color blocks.

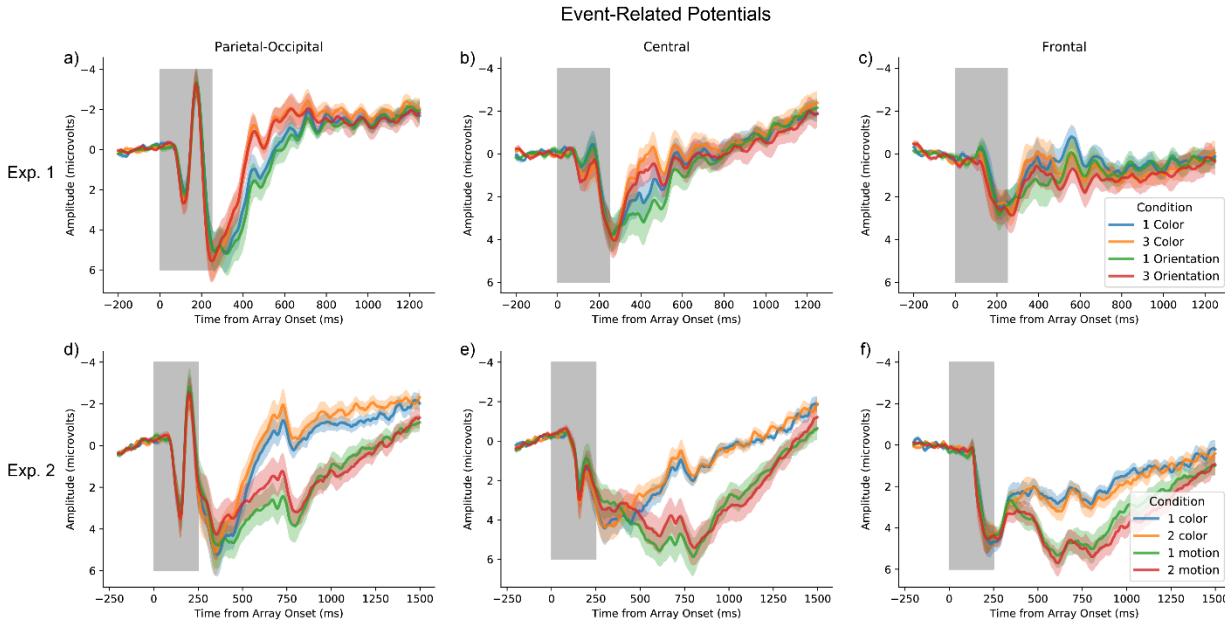

Figure 13 Event-related potentials (ERPs) for each condition in both experiments at parietal/occipital, central, and frontal electrodes. The gray region indicates stimulus presentation.

Significance testing

In all classification analyses, we tested whether classification accuracy was significantly above chance at each time point using a paired-samples, one-tailed t test. Classification accuracy was compared with empirical chance accuracy, defined by testing the trained model on randomly shuffled trial labels. Because we tested for significance at each time

point, we used the Benjamini-Hochberg procedure to control the false-discovery rate (FDR) at .05.

Results

Behavioral

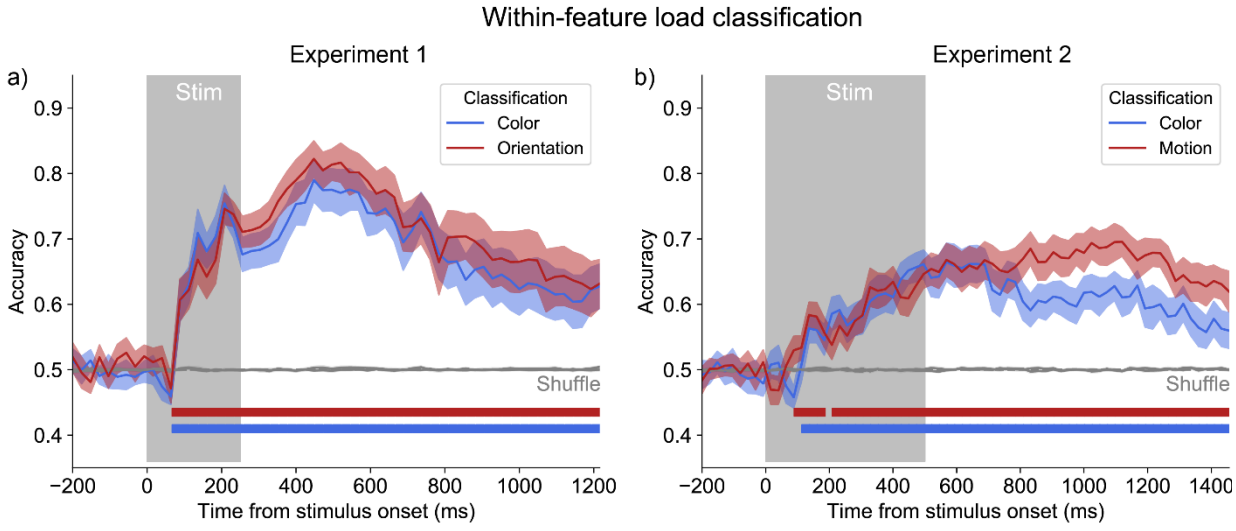

Figure 14 Load classification accuracy for each attended feature over time for (a) Experiment 1 (color and orientation) and (b) Experiment 2 (color and motion coherence). Classification accuracy is indicated with the red and blue lines. The shaded area around the lines indicate the standard error of the mean. Red and blue squares at the bottom indicate time points in which classification accuracy was significantly above chance (corrected $p < .05$, false-discovery rate = .05 with Benjamini-Hochberg procedure). The gray dashed line indicates chance classification accuracy. The vertical gray rectangle indicates the time period during which the memory array was displayed. The shuffle condition reveals empirical chance

accuracy, obtained by training the model on non-permuted data then testing on data with permuted trial labels.

Across both experiments and conditions, subjects performed the task with above-chance accuracy (see Fig. 12). In Experiment 1, a repeated measures analysis of variance (ANOVA) revealed a significant main effect of set size, indicating that accuracy declined as set size increased, $F(1, 11) = 20.59$, $p < .001$. There was also a significant main effect of attended feature, $F(1, 11) = 32.74$, $p < .001$, and a significant interaction of attended feature and set size, $F(1, 11) = 16.72$, $p < .001$. In Experiment 2, a repeated measures analysis of variance (ANOVA) revealed a significant main effect of set size, indicating that accuracy declined as set size increased, $F(1, 12) = 31.47$, $p < .001$. There was also a significant main effect of attended feature, $F(1, 12) = 26.02$, $p < .001$, and a significant interaction of attended feature and set size, $F(1, 12) = 27.76$, $p < .001$.

Within-Feature Load Classification

We successfully predicted WM load. In both experiments (Experiment 1: $n = 13$; Experiment 2: $n = 15$), we used a logistic regression classifier on raw EEG amplitudes from binned trials within subjects (15 trials per bin) at each time bin (25-ms window). We could classify WM load (Experiment 1: Set Size 1 vs. Set Size 3; Experiment 2: Set Size 1 vs. Set Size 2) during the stimulus presentation and throughout the delay period in each attended feature condition (Experiment 1: color and orientation; Experiment 2: color and motion coherence) (Fig. 14; red squares indicate corrected $p < .05$, FDR-controlled at .05 with Benjamini-Hochberg procedure with 51 time bins tested). Above-chance classification was observed starting in early time bins in each experiment (Experiment 1 color: 88-ms to 112-ms time bin; Experiment 1 orientation: 88-ms to 112-ms time bin;

Experiment 2 color: 136-ms to 184-ms time bin; Experiment 2 motion coherence: 112-ms to 136-ms). Classification was sustained throughout the entire delay period for both experiments. Mean classification accuracy (with chance at .5) during the delay period for Experiment 1 color was .688 (SD = .0567), for Experiment 1 orientation mean accuracy was .723 (SD = .611), for Experiment 2 color it was .612 (SD = .029), and for Experiment 2 motion coherence it was .663 (SD = .020).

Across-Feature Load Classification

The main analysis examined whether the load classification generalized across distinct feature values (i.e., color and orientation). In Experiment 1, we trained the classifier using the color trials and tested it on the orientation trials. We also trained on orientation trials and tested on color trials. In Experiment 2, we trained the classifier using the color trials and tested it on the motion coherence trials. We also trained on motion coherence trials and tested on color trials. In Experiment 1, in both directions of training and testing, robust classification was sustained throughout the entire delay period (Fig. 15a-b). Mean classification accuracy during the delay period for color to orientation was .680 (SD = .0593) and for orientation to color was .669 (SD = .0522). In Experiment 2, there was above chance classification accuracy in both directions, however, it was not sustained throughout the entire delay period. Mean classification accuracy during the delay period for color to motion coherence was .551 (SD = .0192) and for motion coherence to color was .538 (SD = .0215). Thus, the same multivariate pattern classified load precisely for memoranda with distinct relevant features, revealing a load-sensitive signal that is separable from the specific content stored in WM. However, in Experiment 2, these

across-feature classification accuracies are significantly lower than those we saw with within-feature classifications for nearly the entire delay period (see Fig. 15c-d).

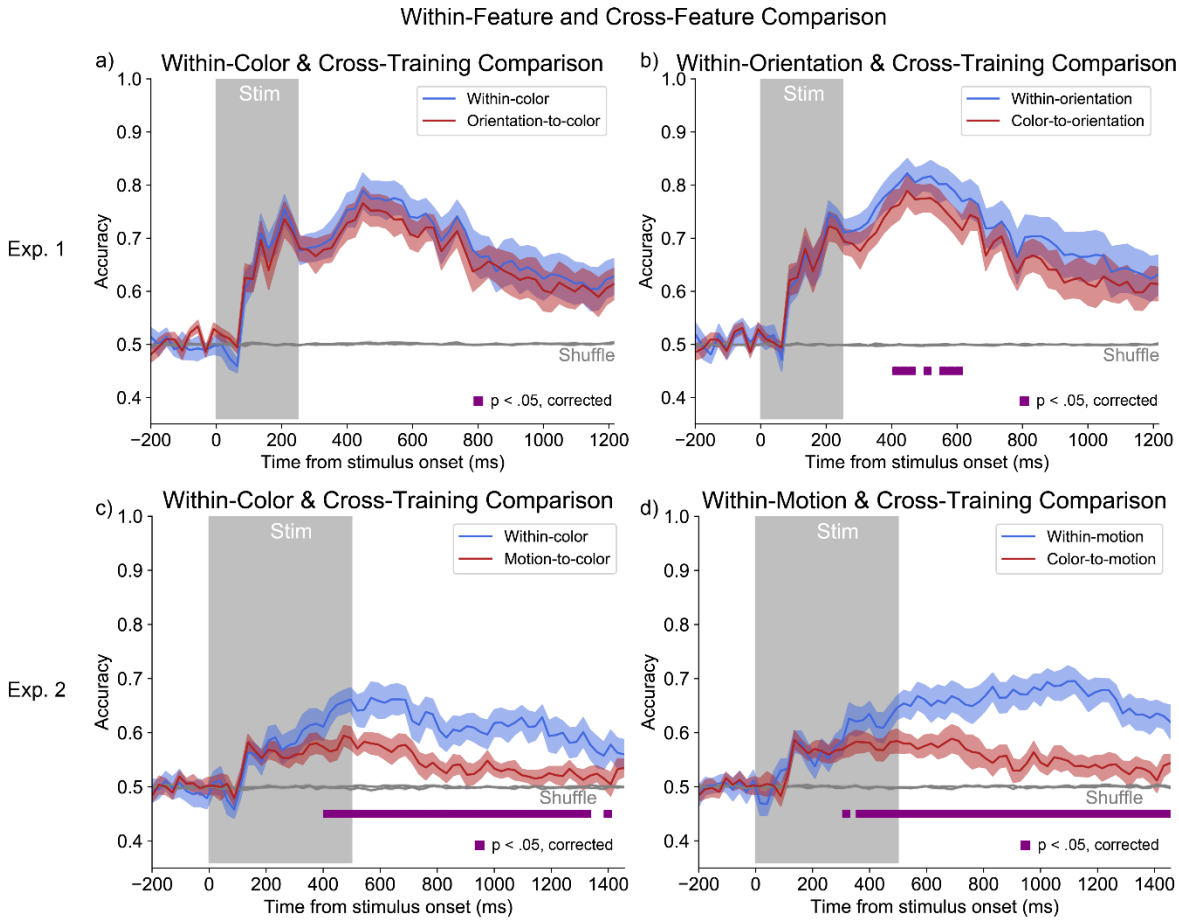

Figure 15 Accuracy for within- and cross-feature load classification. The blue line represents the classification accuracy when trained and tested on data from the same attended feature. The red line represents classification accuracy when trained on data from one feature and tested on the other attended feature. The purple squares represent times when cross-trained classification accuracy is significantly different from within-feature accuracy (corrected $p < .05$, false-discovery rate = .05 with Benjamini-Hochberg procedure). The gray line indicates chance classification accuracy. The gray vertical rectangle indicates the time period during which the memory array was displayed. The shuffle condition reveals empirical chance accuracy, obtained by training the model on non-permuted data then testing on data with permuted trial labels.

Attended Feature Classification

If feature-general classification is the result of EEG lacking the sensitivity to measure differences in color and orientation feature information, then classification of the attended

feature in the task should not be possible. However, in each, we trained the classifier to predict the attended feature (alternated each block). In both experiments, classification was sustained throughout the entire delay period (Fig. 16). In Experiment 1, mean classification accuracy during the delay period was .589 (SD = .298). In Experiment 2, mean classification accuracy during the delay period was .803 (SD = .0678).

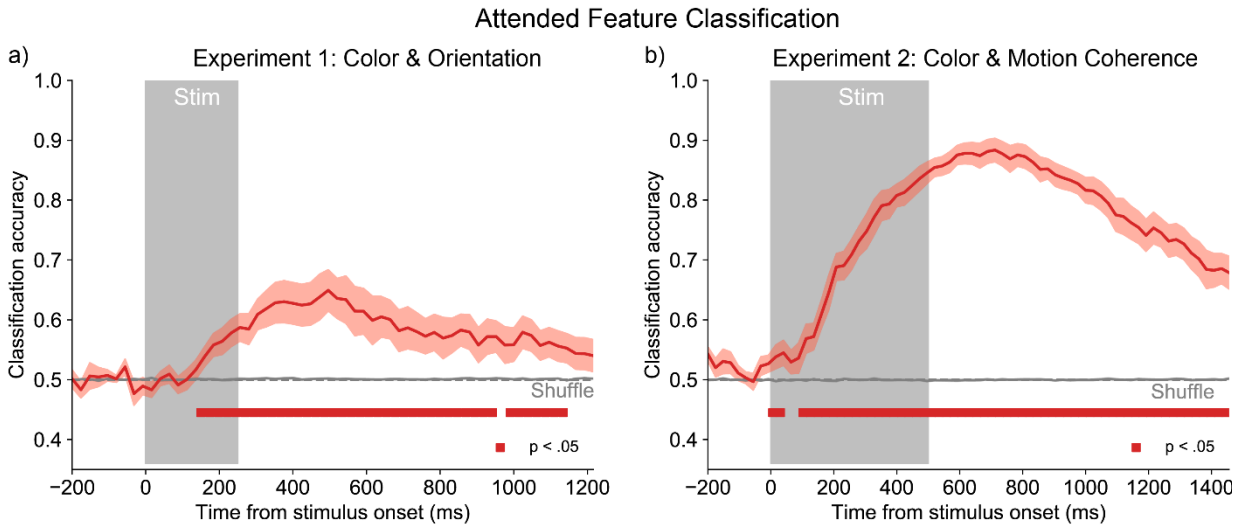

Figure 16 Accuracy for attended feature classification in (a) Experiment 1 and (b) Experiment 2. The red line represents accuracy for predicting which feature the participant is attending to (and subsequently remembering) in a given trial. The red squares represent times when classification accuracy is significantly above chance (corrected p < .05, false-discovery rate = .05 with Benjamini-Hochberg procedure). The gray line indicates chance classification accuracy. The gray vertical rectangle indicates the time period during which the memory array was displayed. The shuffle condition reveals empirical chance accuracy, obtained by training the model on non-permuted data then testing on data with permuted trial labels.

Coherence Level Classification

In addition to the load and attended feature classification, we also found strong classification of coherence level (Fig. 17). Generally, we are unable to classify specific feature

values from EEG (such as color or orientation). Using only set size 1 trials, we classified whether the memory object had coherent or incoherent motion. Mean classification accuracy during the delay period was .566 (SD = .0307). Importantly, classification was at chance for coherence level classification in the attend color blocks (mean = .503, SD = .0208). This provides strong evidence for voluntary storage of features within an object.

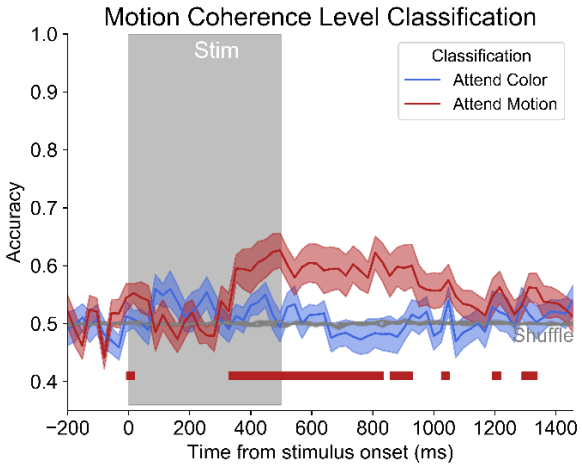

Fig. 17 Classification accuracy for motion coherence. The red line represents accuracy for classifying the motion coherence level of a given Set Size 1 trial while the participant was attending to motion coherence. The blue line presents accuracy for classifying the motion coherence level of a given Set Size 1 trial while the participant was attending to color. The red squares represent times when classification accuracy for the "attend motion" trials is significantly above chance (corrected $p < .05$, false-discovery rate = .05 with Benjamini-Hochberg procedure). There are not time points above chance for the "attend color" trials. The gray line indicates chance classification accuracy. The gray vertical rectangle indicates the time period during which the memory array was displayed. The shuffle condition reveals empirical chance accuracy, obtained by training the model on non-permuted data then testing on data with permuted trial labels.

Discussion

A content-independent pointer system supports on-line storage in working memory. Pointers track items in memory regardless of their features. There are many models of visual WM that include the idea of separable neural processes for content maintenance and the individuation and bindings of those representations (Balaban et al., 2019; Bouchacourt & Buschman, 2019; Oberauer, 2019; Swan & Wyble, 2014; Xu & Chun, 2009). Chapter 1 provides evidence that mvLoad generalizes across items with different featural content. mvLoad generalizes from colors to orientations (and vice versa), as well as from single-feature items to conjunctions items. Problem in different sessions and color/orientation neural interdigitation. However, a weakness of the evidence in

Chapter 1 was that data for items with different feature values came from separate EEG sessions. That could result in weakened cross-training results, thus, underestimating the degree of feature-independence. However, a classifier trained color and orientation items may be particularly well-suited to generalizing to each other because color and orientation neural populations are interdigitated. The use of more neural separable populations such as color and motion may provide a more rigorous test of feature-independence. Here, we provide much stronger evidence for mvLoad's generalization across items with different features values. In 2 experiments, we test the generalizability of mvLoad within single sessions of EEG recording. Experiment 1 uses color and orientation objects and replicates the feature-independence findings of Chapter 1. Experiment 2 uses color and motion coherence objects and provides stronger evidence for the load classifier generalizing across items with unique feature values.

Experiment 1 replicates the cross-training from Chapter 1, where mvLoad was trained on color and tested on orientation and vice versa. This experiment provides more robust evidence because the cross-training is within-session and uses perfectly balanced stimuli. This allows us to perfectly compare within- and across- feature training. We find that training and testing across items with different feature dimensions shows minimal reduction in accuracy as compared to items with the same feature dimension. Importantly, we also show that the attended feature of the item can be classified. This suggests that there is a difference between the two conditions, but the load signal generalizes.

Experiment 2 extends the cross-training to color and motion coherence. This is an important addition, because color and motion neural populations are more disparate than colors and orientations or shapes. It's expected that if cross-training is driven by confusion between the feature-specific neural activity, then load decoding should drop to chance. While we do show evidence of a drop in accuracy during cross-training (as compared to within-feature load classification), there is still sustained above-chance load classification. This suggests that while there is clearly some feature-specific information present, there is still a robust underlying signature of load that generalizes across very distinct features.

There is a variety of evidence that is in line with the content-independent pointer system. If WM storage is limited by pointer deployment, this could explain why there is a strong correlation between fluid intelligence and WM capacity but not precision (Fukuda et al., 2010). Also, many studies have shown an object-based benefit in which more features can be maintained in multi-feature objects than single-feature objects (Luck & Vogel, 1997; Olson & Jiang, 2002). There is also evidence that different regions of visual cortex track the number and complexity of objects in WM (Xu & Chun, 2009).

In summary, we presented further evidence for a content-independent load signature in raw EEG topography. It generalizes across colors to orientations with almost no loss in classification accuracy. It also generalizes across colors and motion coherence, which are represented by very disparate neural populations. This strengthens the argument that the generalization isn't just the result of interdigitated neural populations resulting in similar EEG activity. In addition, we show strong classification of attended feature which

further suggests that the EEG activity is distinct for different remembered features, but the load classification still generalizes. This evidence is in line with the hypothesis that mvLoad is tracking content-independent pointers. The pointer system supports the individuation and maintenance of objects over time and space, while being separate from the maintenance of the feature values of the items themselves.

Chapter 3: Spatial attention and working memory gating are distinct forms of voluntary attentional control

Introduction

Past work has revealed a tight relationship between storage in visual working memory (WM) and covert spatial attention (Awh et al., 2006; Awh & Jonides, 2001). One of the most compelling forms of evidence comes from studies that have examined spatially selective alpha oscillations – known to track the precise locus of covert spatial attention (Foster, Sutterer, et al., 2017; Rihs et al., 2007; Samaha et al., 2016) – during the delay period of working memory tasks. For example, Foster et al. (2016) showed that the scalp topography of alpha activity precisely tracks positions that are maintained in working memory, and that this signal is sustained throughout the blank delay period of the WM task (Foster et al., 2016). Moreover, the same signature of spatial attention is observed during WM storage even when location is completely irrelevant to the task (Foster, Bsales, et al., 2017). Finally, WM performance declines when observers are prevented from deploying visual attention towards memorized locations (Awh et al., 1998; Williams et al., 2013), suggesting that attention plays a functional role in WM storage. These findings notwithstanding, a fundamental question remains regarding the relationship between spatial selective attention and visual working memory. Is focusing spatial attention on an item tantamount to encoding it into working memory, or are these processes separable but intertwined aspects of attentional control?

Indeed, there is an emerging body of evidence that indicates a potential dissociation of spatial selective attention and WM storage. For example, Hakim et al. (2019) examined the EEG correlates of spatial attention and WM storage in two distinct conditions (Hakim et al., 2019). When observers were instructed to storage an initial array of items in working memory, both lateralized alpha (tracking spatial attention) and contralateral delay activity (CDA, known to track the number of items stored in visual WM) tracked the locus and number of items in the sample array. However, when observers were presented with identical stimulus displays and instructed to covertly attend the marked positions, CDA activity was virtually eliminated while lateralized alpha activity continued to track the attended hemifield. Likewise, Hakim et al. (2021) found that while CDA activity was eliminated for irrelevant items presented on one side of space, lateralized alpha signals were insensitive to target relevance (Hakim et al., 2021). These findings suggest that spatial attention and WM storage may be intertwined but separable aspects of voluntary attentional control. One limitation of these findings, however, is that lateralized alpha activity is a relatively coarse measure of covert spatial orienting that does not determine whether spatial attention is precisely oriented towards a given item, or merely directed towards the visual field containing that item. This limitation prevents a clear conclusion about whether spatially attending a specific item entails the entry of that item into working memory.

Thus, in the present work, we employed an inverted encoding model (IEM) that enabled us to track the deployment of spatial attention towards specific items (Foster, Sutterer, et al., 2017). In addition we used a recently developed multivariate approach (mvLoad) for

decoding the number of discrete items encoded into visual WM (Adam et al., 2020; Thyer et al., 2022). Another innovation in our design is that targets and distractors were randomly positioned such that observers could not direct attention in advance towards specific locations. Thus, we were able to observe the goal-driven encoding of targets into working memory, while simultaneously measuring which elements were spatially attended. To anticipate the results, mvLoad analyses revealed that observers could voluntarily gate encoding into working memory, such that targets could be selectively stored in visual working memory while distractors were held out. By contrast, alpha oscillations revealed that spatial selective attention was precisely oriented towards both elements in the sample display, and that the sharpness of this spatial spotlight did not depend on target status. In other words, at precisely the same moments in time where mvLoad analyses revealed selective encoding of targets into working memory, alpha oscillations showed that spatial attention was sharply

Methods

Subjects

The experiment included n volunteers participating for monetary compensation ($15-20 per hr). Subjects were between the ages of 18 and 35 years old, reported normal or correct-to-normal visual acuity, and provided informed consent according to procedures approved by The University of Chicago Institutional Review Board. Subjects were recruited via online advertisements and fliers posted on the university campus.

Experiment 1. 31 volunteers participated in Experiment 1 (20 females; mean age = 25.7 years, SD = 3.9). 8 subjects were excluded from the final sample for the following reasons: The session was ended early due to eye movements (n = 3); the subject's data was corrupted or otherwise unusable (n = 5). The final sample size was 23 (16 female; mean age = 26.0 years, SD = 4.12).

Experiment 2. 23 volunteers participated in Experiment 2 (16 females; mean age = 26.2 years, SD = 3.8). 3 subjects were excluded from the final sample for the following reasons: the subject didn't have enough data after artifact rejection (n = 2); the subject's data was corrupted or otherwise unusable (n = 1). The final sample size was 20 (15 female; mean age = 26.4 years, SD = 3.7).

Apparatus

We tested the subjects in a dimly lit, electrically shielded chamber. Stimuli were generated using PsychoPy (Peirce et al., 2019). Subjects viewed the stimuli on a gamma-corrected 24-in. LCD monitor (refresh rate = 120 Hz, resolution = 1,080 × 1,920 pixels) with their chins on a padded chin rest at a viewing distance of 75 cm.

Memory displays

Stimuli were presented against a mid-gray background (~61 cd/m^2). Memory arrays included one or two to-be remembered items. In experiment 1, distractor items were either letters (A, E, G, J, M, P, T, or X) or hashtags (#). In experiment 2, distractors were rectangles of equal area to the target squares. They were either shown using the same

set of colors as memory items, or shown in a shade of gray (red, green, blue [RGB] value = 163, 163, 163) that matched the average luminance of possible colors in the color set.

Task procedures

Experiment 1. The experiment used a whole-field change detection task (Fig. 18a). On each trial, a memory array appeared containing 1 or 2 items. Items were either digits, letters, or hashtags (#). There were 4 experimental conditions. In the single target condition, one digit would appear as a memory item. In the single target plus hashtag condition, one digit would appear as a memory item and one hashtag would appear as a distractor. In the single target plus letter distractor condition, one digit would appear as a memory item and one letter would appear as a distractor. In

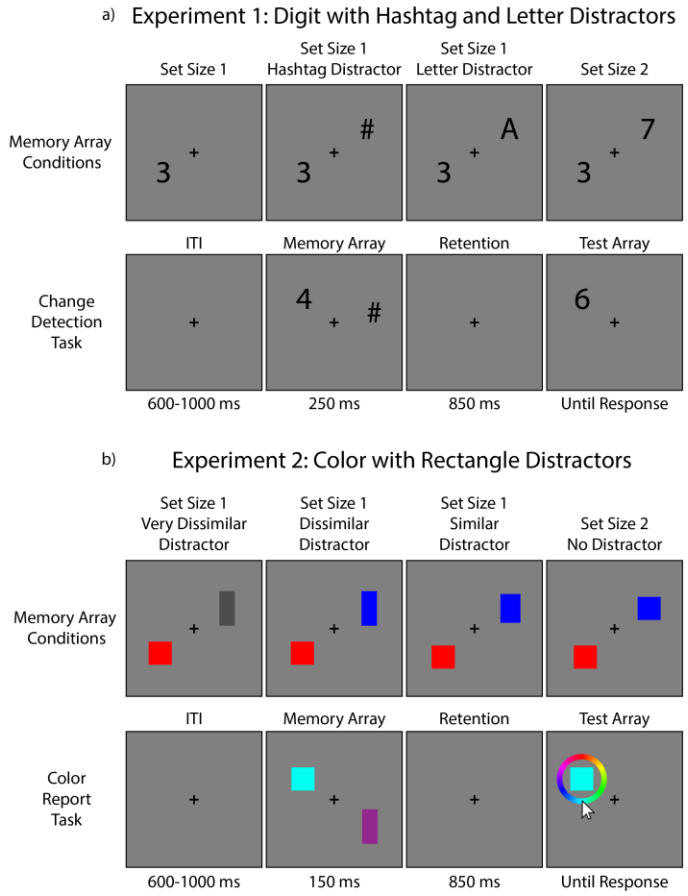

Figure 18 Task schematics for an example Set Size 3 trial in the whole-field change-detection task used on both experiments. In Experiment 1 (a), subjects were cued at the beginning of each block to either remember the color or orientation of the memory items while ignoring the gray distractor items. In Experiment 2 (b), subjects were cued at the beginning of each block to either remember the color or the motion coherence of the memory items while ignoring the gray distractor items. During change trials, the attended feature changes 50% of the time.

the two target condition, two digits would appear as memory items. Conditions were blocked. Items were positioned with one item per quadrant. Two memory items never appeared in one quadrant together, and all items were placed at least 4° degrees apart. Subjects viewed a memory array (200 ms), remembered items across a delay (800 ms), were probed on one memory item, and reported whether the probed item was the same as or different from the remembered item (unspeeded).

Experiment 2. The experiment used a whole-field change detection task (Fig. 18b). On each trial, a memory array appeared containing 1 or 2 items. Memory items were colored squares (dimensions 1° by 1°) and distractor items were colored or gray rectangles. There were 4 experimental conditions. In the very dissimilar distractor condition, 1 colored square plus 1 gray rectangle (dimensions .55° by 1.82°) appeared. In the dissimilar distractor condition, 1 colored square plus 1 colored rectangle (dimensions .55° by 1.82°) appeared. In the similar distractor condition, 1 colored square plus 1 colored rectangle (dimensions .67° by 1.50°) appeared. Conditions were blocked. Items were placed 4° from fixation, within one of five bins (54°-126°, 126°-198°, 198°-270°, 270°-342°, 342°-54°). These bins were later used to group items for *Inverted Encoding Modeling*. Items were placed within the bins with counterbalancing, such that every combination of item placement was equally probable. Items were placed at least 2° (degrees of visual angle) apart. Subjects viewed a memory array (150 ms), remembered items across a delay (850 ms), were probed on one memory item, and reported the color of the probed item by clicking on a color wheel (unspeeded). The probed item was indicated by the location of the color wheel.

EEG acquisition

We recorded EEG activity from 30 active Ag/AgCl electrodes mounted in an elastic cap (Brain Products actiCHamp, Munich, Germany). We recorded from international 10-20 sites Fp1, Fp2, F7, F3, Fz, F4, F8, FT9, FC5, FC1, FC2, FC6, FT10, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, and O2. Two additional electrodes were affixed with stickers to the left and right mastoids, and a ground electrode was placed in the elastic cap at position Fpz. All sites were recorded with a right-mastoid reference and were rereferenced off-line to the algebraic average of the left and right mastoids. Data were filtered on-line (low cut-off = 0.01 Hz, high cut-off = 80 Hz, slope from low to high cut-off = 12 dB/octave) and were digitized at 500 Hz using BrainVision Recorder (Brain Products, Munich, Germany) running on a PC. Impedance values were brought below 10 kΩ at the beginning of the session.

Eye tracking

We monitored gaze position using a desk-mounted EyeLink 1000 Plus infrared eye-tracking camera (SR Research, Ontario, Canada). Gaze position was sampled at 1000 Hz. According to the manufacturer, this system provides spatial resolution of .01° of visual angle and average accuracy of 0.25 to 0.50° of visual angle. We calibrated the eye tracker every one to two blocks of the task and between trials during the blocks if necessary. We drift-corrected every 5 or 10 trials. Additionally, we drift-corrected the eye-tracking data for each trial by subtracting the mean gaze position measured during a 200-ms window immediately before the onset of the memory array.

Artifact rejection

We segmented the EEG data into epochs time-locked to the onset of the memory array (200 ms before until 1,000 ms after stimulus onset). We baseline-corrected the EEG data by subtracting mean voltage during the 200-ms window immediately prior to stimulus onset. Eye movements, blinks, blocking, drift, and muscle artifacts were detected by applying automatic criteria and we discarded any epochs contaminated by artifacts.

Eye movements and blinks

We employed real-time eye movement detection. If participants moved their eyes more than 1.25° from fixation, the trial was interrupted and their eye position was shown to them for feedback purposes. Interrupted trials were made up at the end of the block. During preprocessing, we rejected trials that contained eye movements beyond 1° of visual angle using the pop_artextval function in ERPLAB (Lopez-Calderon & Luck, 2014).

Drift, muscle artifacts, and blocking

We checked for drift (e.g., skin potentials) with the pop_rejtrend function in ERPLAB. We checked for muscle artifacts with the pop_artextval function in ERPLAB. We excluded trials where EEG activity was greater than 80 μV or less than -80 μV. We excluded trials with peak-to-peak activity greater than 100 μV within a 200-ms window with 100- ms steps. We also excluded trials with any value beyond a threshold of 80 μV.

mvLoad procedure

Load classification

The mvLoad analysis is within-subjects classification of WM load on baselined EEG. Although our approach allowed robust above-chance performance with single trials, we used randomly chosen groups of 20 trials within each set size and attended feature to increase signal-to-noise ratio. We divided each trial into 50-ms windows with 25-ms steps and calculated the average voltage for each electrode in the window. Classification was performed using the Scikit-Learn Logistic Regression model. The classifier was trained to discriminate between conditions. Classification was tested on a held-out set of data using the train_test_split function from Scikit-Learn (Pedregosa et al., 2011) stratified on the trial conditions. This cross-validation procedure splits the data in 80% training and 20% testing sets while preserving the percentage of samples for each load condition. This split was repeated 1,000 times, and results for each subject and time point were averaged across these repetitions. Training data were standardized at each time point using the StandardScaler Scikit-Learn function, and test data were standardized using the mean and standard deviation of the training set.

Significance testing

In all classification analyses, we tested whether classification accuracy was significantly above chance at each time point using a paired-samples, one-tailed t test. Classification accuracy was compared with empirical chance accuracy, defined by testing the trained model on randomly shuffled trial labels. Because we tested for significance at each time point, we used the Benjamini-Hochberg procedure to control the false-discovery rate (FDR) at .05.

Hyperplane procedure

In order to further leverage classification, we can use the average classification confidence score of a given condition after training the classifier. The confidence score comes from the decision_function method. This score is proportional to the signed distance of that sample to the hyperplane. By training a binary classifier, and testing on other conditions, we can track the average classification confidence level for each condition over time. For training and testing, we use the same procedures as typical classification, but we remove certain conditions from the training set, so they only appear in the testing set. This allows us to set “prototypical” classes for the classifier by only training on them, and then measure how other conditions are classified by testing on them.

Inverted Encoding Models

We employed an Inverted Encoding Model to track spatially-specific alpha-activity in the EEG signal (Brouwer & Heeger, 2009; Sprague & Serences, 2013). This method allows us to estimate the responses of spatial channels (CTFs) which reflect the spatial tuning of population-level alpha power over time. Thus, we have a temporally resolved method to track covert spatial attention (Foster, Bsales, et al., 2017; Foster et al., 2016).Crucially, we can use this technique to compare the deployment of spatial attention to both targets and distractors. In Experiment 1, the location bins were set at each quadrant because we did not initially intend to use IEMs so the stimulus placement is not optimized for them. In

Experiment 2, we positioned stimuli within 5 different bins around fixation (see Task Procedures - Experiment 2 for details).

Results

Behavioral

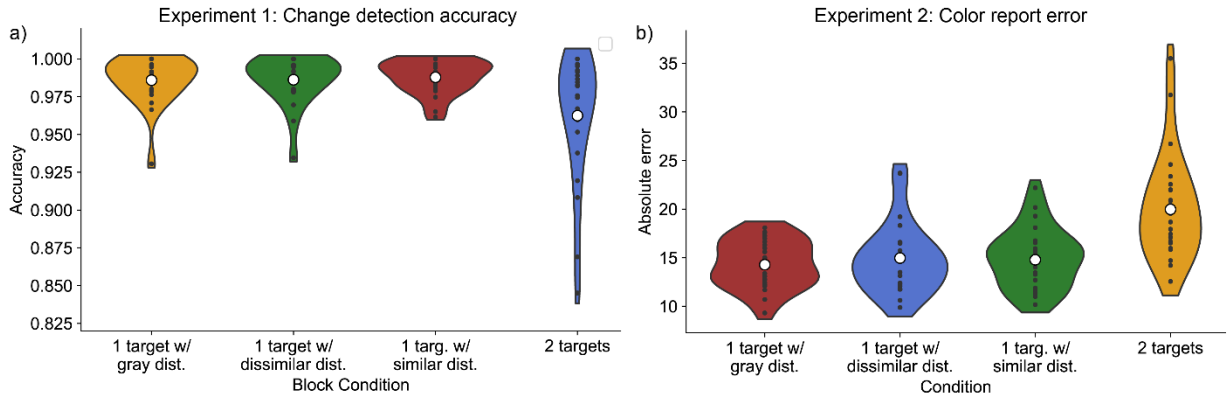

Figure 19 Task accuracy for each condition in both experiments. Experiment 1 (a) change detection accuracy and Experiment 2 (b) color report absolute error. Black dots indicate individual data, white dots indicate means, and shaded regions indicate the density of the data.

Across both experiments and all conditions, subjects performed the task with above chance accuracy (see Fig. 19) In Experiment 1, change detection accuracy ranged from 0.96-0.99. In Experiment 1, all set size 1 condition accuracies were not significantly different from each other. 1 target (M = .99, SD = .12) was not significantly different from 1 target plus a hashtag distractor (M = .99, SD = .12), $t(21) = -.129$, $p = .900$, or 1 target plus a letter distractor (M = .99, SD = .11), $t(21) = -.860$, $p = .401$. 1 target plus a hashtag distractor was not significantly different from 1 target plus a letter distractor, $t(21) = -.760$, $p = .456$. After averaging set size 1 condition accuracies (M = .99, SD = .01), set size 1 accuracy was significantly different from set size 2 accuracy (M = .96, SD = .19), $t(21) = 3.33$, $p < .05$.

Experiment 2, color report absolute error ranged from 14.3-20.0 degrees. 1 target plus a gray distractor (M = 14.28, SD = 13.00) was not significantly different from 1 target plus a dissimilar distractor (M = 14.78, SD = 15.26), t(23) = -.98, p = .339, or 1 target plus a similar distractor (M = 14.96, SD = 15.10), t(23) = -1.46, p = .157. 1 target plus a dissimilar distractor was not significantly different from 1 target plus a similar distractor, t(23) = -.27, p = .789. After averaging set size 1 condition accuracies (M = 14.73, SD = 2.78), set size 1 accuracy was significantly different from set size 2 accuracy (M = 19.97, SD = 24.67), t(23) = -6.29, p < .001.

Hyperplane analysis

Figure 19a (Experiment 1) and Figure 20a (Experiment 2) illustrate the output over time of a classifier that was trained exclusively on 1 target and 2 target trials and then tested on all conditions. The output here is from the classifier's decision_function method, which returns the confidence score of the sample. This score is proportional to the signed distance of that sample to the hyperplane. Stronger evidence for Set Size 2 is plotted in the positive direction, whereas stronger evidence for Set Size 1 is plotted in the negative direction.

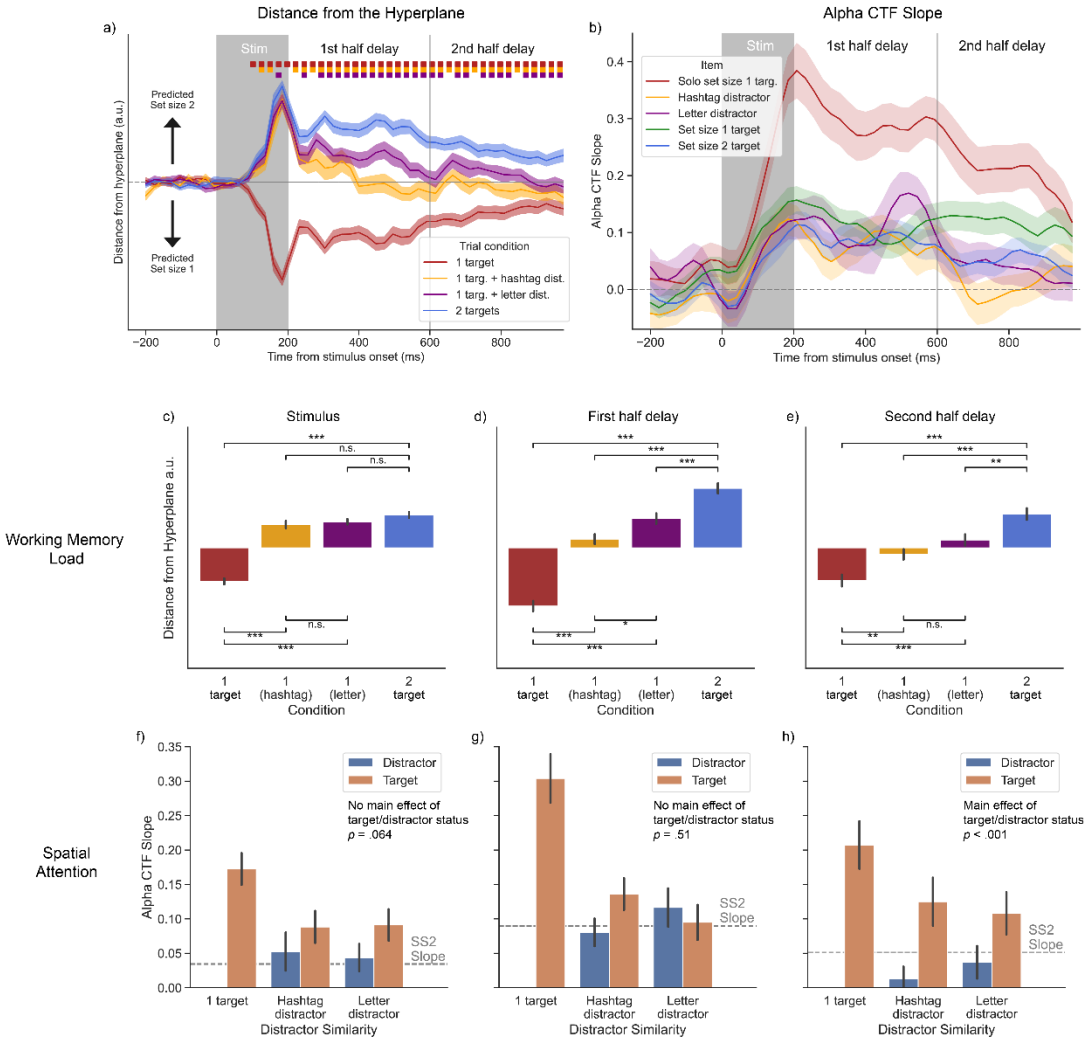

Figure 20 Hyperplane and CTF results from Experiment 1. a) The distance from the hyperplane for each condition over time. After training the classifier on Set Size 1 and 2, we tested it on all conditions (see Hyperplane analysis section for more detail). Each line represents the distance from the hyperplane over time for that condition. The colored shaded region is the standard error of the mean. The colored squares above indicate time points in which the matching line is significantly different from Set Size 2. b) CTF selectivity (as measured by slope) over time for each condition. The shaded region is the standard error of the mean. c-e) The average distance from the hyperplane for each condition during stimulus presentation (c), first half of the delay period (d), and second half of the delay period (e). f-h) The average CTF slope for the target and distractor for each condition during the stimulus presentation (f), first half of the delay period (g), and second half of the delay period (h).

a)

COMPARISON	STIMULUS	1ST HALF DELAY	2ND HALF DELAY
1 TARGET – HASHTAG DIST.	T = -1.42, p < .001	T = -1.67, p < .001	T = -.65, p < .01
1 TARGET – LETTER DIST.	T = -1.48, p < .001	T = -2.18, p < .001	T = -1.01, p < .001
1 TARGET – 2 TARGET	T = -1.66, p < .001	T = -2.93, p < .001	T = -1.66, p < .001
HASHTAG DIST. – LETTER DIST.	T = -.066, p = .92	T = -.51, p < .05	T = -.36, p = .27
HASHTAG DIST. – 2 TARGET	T = -.24, p = .11	T = -1.26, p < .001	T = -1.01, p < .001
LETTER DIST. – 2 TARGET	T = -.17, p = .36	T = -.75, p < .001	T = -.65, p < .01

b)

CONDITION	STIM	1ST HALF DELAY	2ND HALF DELAY
1 TARGET	M = -.84, SD = .33	M = -1.45, SD = .60	M = -.81, SD = .68
HASHTAG DISTRACTOR	M = .58, SD = .42	M = .22, SD = .57	M = .-.16, SD = .60
LETTER DISTRACTOR	M = .65, SD = .31	M = .73, SD = .61	M = .20, SD = .65
2 TARGET	M = .82, SD = .32	M = 1.48, SD = .59	M = .84, SD = .67

Table 1 Results from the time window hyperplane analysis for Experiment 1 (see Fig Xc-e). a) The results of the Tukey HSD test of each pairwise comparison for the experiment conditions distance from the hyperplane. b) Mean and standard deviation for each condition in during each time window.

Experiment 1. The classifier was trained to classify 1 target versus 2 target trials. The trained classifier was then tested on all conditions. The 1 target and 1 target with a hashtag distractor were classified significantly differently from Set Size 2 throughout the delay period starting during stimulus presentation (1 target: 112ms-136ms time bin, 1 target with hashtag distractor: 136ms-160ms time bin) (Fig. 20a). The 1 target with a letter distractor condition was not as consistently significantly different from the 2 target condition, but a generally sustained significant difference started during the first half of the delay period (304ms-328ms time bin). Thus, the mvLoad analysis shows a clear distinction between Set Size 2 and the Set Size 1 conditions with a distractor. To further highlight the trajectory of the hyperplane, we separate the time course into 3 bins: stimulus

display (0 - 200 ms), first-half of the delay period (200 - 600 ms), and second half of the delay period (600-1000 ms) (Fig. 20c-e). We compared each condition using a Tukey HSD test. During stimulus presentation, the 1 target with a hashtag distractor

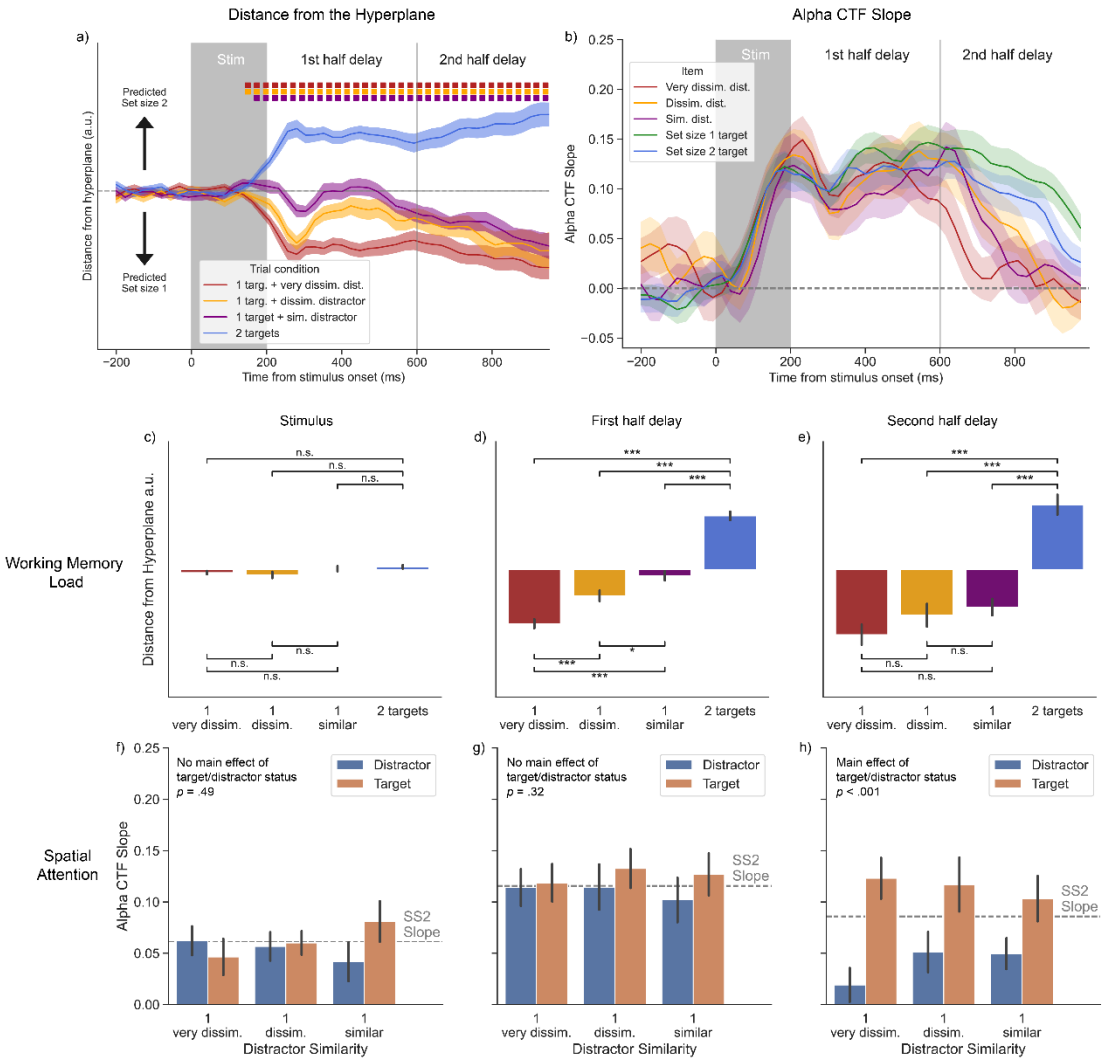

Figure 21 Hyperplane and CTF results from Experiment 2. a) The distance from the hyperplane for each condition over time. After training the classifier on Set Size 1 and 2, we tested it on all conditions (see Hyperplane analysis section for more detail). Each line represents the distance from the hyperplane over time for that condition. The colored shaded region is the standard error of the mean. The colored squares above indicate time points in which the matching line is significantly different from Set Size 2. b) CTF selectivity (as measured by slope) over time for each condition. The shaded region is the standard error of the mean. c-e) The average distance from the hyperplane for each condition during stimulus presentation (c), first half of the delay period (d), and second half of the delay period (e). f-h) The average CTF slope for the target and distractor for each condition during the stimulus presentation (f), first half of the delay period (g), and second half of the delay period (h).

and with a letter distractor were not significantly different from the 2 target condition, and were significantly different from the 1 target condition. During the first half of the delay period, all pairs were significantly different from each other. Importantly, all Set Size 1 conditions were significantly different from Set Size 2. Additionally, distance from the hyperplane monotonically increased with distractor difficulty. During the second half of the delay period, each Set Size 1 condition was still significantly different from the Set Size 2 condition. (See table 1 for values)

Experiment 2. The classifier was trained to classify Set Size 1 trials with a gray distractor and Set Size 2 trials (Fig. 21a). The trained classifier was then tested on all conditions. For Set Size 1 conditions, all three conditions were classified significantly differently from Set Size 2 throughout the whole delay period starting during the delay period (Very dissimilar distractor: 160ms-184ms time bin; Dissimilar distractor: 160ms-184ms time bin; Similar distractor: 184ms-208ms time bin). Thus, the mvLoad analysis shows a clear distinction between Set Size 2 and the Set Size 1 conditions with a distractor. To further highlight the trajectory of the hyperplane, we separate the time course into 3 bins: stimulus display (0 - 200 ms), first-half of the delay period (200 - 600 ms), and second half of the delay period (600-1000 ms) (Fig. 21c-e). We compared each condition using a Tukey HSD test. During stimulus presentation, hyperplane distance was not significantly different for any given pair of conditions. During the first half of the delay period, all pairs were significantly different from each other. Importantly, all Set Size 1 conditions were significantly different from Set Size 2. Additionally, distance from the hyperplane monotonically increased with distractor difficulty. During the second half of the delay

period, each Set Size 1 condition was still significantly different from the Set Size 2 condition, however the distractor difficulty effect is no longer present. (See table 2 for values)

IEM analysis

In both experiments, we tested the hypothesis that distractors and target alike were both spatially attended but only targets were stored in WM. We tested this by comparing the slope of the channel tuning functions (CTF) for the Set Size 1 & 2 targets versus the Set Size 1 distractors. If the CTF selectivity (as measured by the slope) is equal for both targets and distractors, then that provides evidence that both are equally spatially attended. And if that selectivity is equal while mvLoad analyses show that the distractors are not in WM, then that shows that items can be attended while not being gated into WM.

a)

COMPARISON	STIMULUS	1ST HALF DELAY	2ND HALF DELAY
V. DISS. DIST. – DISS. DIST.	T = .055, p = .90	T = -.65, p < .001	T = -.46, p = .52
V. DISS. DIST. – SIM. DIST.	T = -.091, p = .65	T = -1.13, p < .001	T = -.65, p = .21
V. DISS. DIST. – 2 TARGET	T = -.14, p = .32	T = -2.53, p < .001	T = -3.05, p < .001
DISS. DIST. – SIM. DIST.	T = -.15, p = .25	T = -.47, p < .05	T = -.19, p = .94
DISS. DIST. – 2 TARGET	T = -.19, p = .08	T = -1.89, p < .001	T = -2.60, p < .001
SIM. DIST. – 2 TARGET	T = -.044, p = .94	T = -1.41, p < .001	T = -2.41, p < .001

b)

CONDITION	STIM	1ST HALF DELAY	2ND HALF DELAY
VERY DISSIM. DISTRACTOR	M = -.13, SD = .36	M = -.62, SD = .61	M = -1.08, SD = 1.31
DISSIMILAR DISTRACTOR	M = .017, SD = .29	M = -.15, SD = .53	M = -.88, SD = .95
SIMILAR DISTRACTOR	M = -.074, SD = .19	M = -1.28, SD = .52	M = -1.53, SD = 1.16
2 TARGET	M = .061, SD = .21	M = 1.26, SD = .49	M = 1.52, SD = 1.15

Table 2 Results from the time window hyperplane analysis for Experiment 2 (see Fig Xc-e). a) The results of the Tukey HSD test of each pairwise comparison for the experiment conditions distance from the hyperplane. b) Mean and standard deviation for each condition in during each time window.

Experiment 1. We found evidence of sustained CTF selectivity for both Set Size 1 distractors and Set Size 1 and 2 targets. As expected, the solo Set Size 1 target had the largest selectivity because it was presented alone. To further highlight the trajectory of the hyperplane, we separate the time course into 3 bins: stimulus display (0 - 200 ms), first-half of the delay period (200 - 600 ms), and second half of the delay period (600-1000 ms). We ran a two-way analysis of variance (ANOVA) on CTF slope within each time window. During the stimulus presentation, there was a significant main effect of condition, $F(3, 126) = 7.46$, $p < .001$, but no significant main effect of target/distractor status, $F(1, 126) = 3.48$, $p = .064$, nor a significant interaction of condition and target/distractor status, $F(3, 126) = .83$, $p = .48$. During the first half of the delay period, there was a significant main effect of condition, $F(3, 126)$ 16.89, $p < .001$, but no significant

main effect of target/distractor status, F(1, 126) = .0064, p = .51, nor a significant interaction of condition and item, F(3, 126) = .92, p = .43. During the second half of the delay period, there was a significant main effect of condition, F(3, 126) = .7.49, p < .001, a significant main effect of target/distractor status, F(1, 126) = 11.26, p = .001, but there was no significant interaction of condition and item, F(3, 126) = .34, p =

.79. Importantly, during the first half of the delay period, mvLoad analyses show a clear load difference in these conditions (see decoding results). That is evidence for distractors being spatially attended, while simultaneously being gated from WM. However, we recognize that Experiment 1 is not perfectly suited for these analyses. Due to the stimulus imbalance of the solo set size 1 target compared to the other conditions, the early load classification is strongly biased towards predicting set size 2 for the hashtag and letter distractor conditions. Additionally, the distractor difficulty is not clearly monotonic. We made an assumption that the hashtag would be easier to filter than the letter distractor. And finally, we didn't originally intend to do IEM analyses on this experiment, so the stimulus presentation was not well-suited for location binning. We binned the items by quadrant which is fairly course compared to the Experiment 2 location binning (see IEM analysis section for details).

Experiment 2. We found evidence of sustained CTF selectivity for both Set Size 1 distractors and Set Size 1 and 2 targets. To further highlight the trajectory of the hyperplane, we separate the time course into 3 bins: stimulus display (0 - 200 ms), first-half of the delay period (200 - 600 ms), and second half of the delay period (600-1000

ms). We ran a two-way analysis of variance (ANOVA) on CTF slope within each time window. During the stimulus presentation, there was no significant main effect of condition, $F(3, 161) = .077$, $p = .97$, item, $F(1, 161) = .48$, $p = .49$, or a significant interaction of condition and item, $F(3, 161) = 1.38$, $p = .25$. During the first half of the delay period, there was no significant main effect of condition, $F(3, 161) = .086$, $p = .97$, item, $F(1, 161) = 1.01$, $p = .32$, or a significant interaction of condition and item, $F(3, 161) = .34$, $p = .79$. During the second half of the delay period, there was no significant main effect of condition, $F(3, 161) = .20$, $p = .90$, however there was significant main effect of item, $F(1, 161) = 21.07$, $p < .005$, but there was no significant interaction of condition and item, $F(3, 161) = .69$, $p = .56$. Importantly, during the first half of the delay period, mvLoad analyses show a clear load difference in these conditions (see decoding results). That is evidence for distractors being spatially attended, while simultaneously being gated from WM.

Discussion

Late selection theory posits that there are multiple stages of selection. There is an early stage where simple stimulus characteristics, form, and meaning can be extracted in parallel (Duncan, 1980). And then there is a later capacity-limited system that is responsible for consolidation in WM (Shiffrin & Schneider, 1977). Similarly, Chun and Potter 1995 proposed a two-stage model of attention which consists of an early parallel system that identifies potential targets and a later system responsible for more detailed processing and conscious perception (Chun & Potter, 1995). These frameworks suggest that information can be attended to and identified as a target before it is gated into WM.

Neural evidence provides additional evidence for a dissociation between early and late processing. Words presented during the attentional blink, while not entering conscious awareness, are processed to the point of meaning extraction as evidenced by an elicited n400 (Luck et al., 1996) and P1 and N1 components. However, there is no P3 component which typically reflects the updating of working memory (Vogel et al., 1998). Additional evidence that uses alpha power topography in EEG suggests that allocation of spatial attention and WM storage are distinct processes. A spatial attention and spatial WM task both produced sustained lateralized alpha power (which indicates deployed spatial attention), however only the WM task produced a CDA (which indicates WM storage) (Hakim et al., 2019). Gunseli 2019 shows evidence that with highly reliable retro-cues in a WM task, uncued items are unattended and dropped from WM (as indexed by lateralized alpha and CDA) (Günseli et al., 2019). However, with unreliable retro-cues, uncued items are unattended but not dropped from WM. Wang et al 2021 proposes that alpha indexes distractor suppression and the extent of spatial attention while being distinct from actual object storage in WM.

If spatial attention and WM gating reflect distinct processes, then distractors can be attended to, identified as distractors, and ultimately not assigned a pointer. To test this hypothesis, we use the *mvLoad* and *inverted encoding model* (IEM) analyses to attempt to dissociate the processes of spatially attending objects and assigning working memory pointers. The mvLoad analysis enables the precise tracking of items individuated in working memory and, we propose, reflects the capacity-limited pointer system that is

responsible for on-line storage of information (Thyer et al., 2022). Using an IEM with total alpha topography provides us with a tool to track the allocation of covert spatial attention (Foster et al., 2016). Both of these analyses have a high time resolution which enables a clear test of whether orienting of spatial attention is indeed yoked to the entry of items into working memory *at the same moments in time*. If attended items are necessarily in WM, then the time course of the IEM and mvLoad analyses will reflect that as soon as items are spatially attended they are loaded into working memory. And similarly, if items are dropped from spatial attention, then they are no longer held in WM.

We provide evidence that items can be covertly attended to, processed, and filtered, without a pointer being assigned and, thus, not stored in working memory. In 2 change detection experiments, we asked participants to remember targets and ignore distractors. In Experiment 1, they we instructed to remember digits and ignore letters and hashtags (#). In Experiment 2, they were presented with colored squares to remember and colored and gray rectangles to ignore. In both experiments, we provide evidence that subjects spatially attended to both distractors and memory items as measured by alpha CTFs. However, we also show that at the same moments where participants are attending these distractors, they are also successfully gating them from WM. This supports the hypothesis that spatial attention and WM selection are distinct but related processes. It also strengthens the argument that WM pointers are not solely individuated using spatial location.

Bibliography

Adam, K. C. S., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The Contribution of Attentional Lapses to Individual Differences in Visual Working Memory Capacity. *Journal of Cognitive Neuroscience*, *27*(8), 1601–1616. https://doi.org/10.1162/jocn_a_00811

Adam, K. C. S., Vogel, E. K., & Awh, E. (2020). Multivariate analysis reveals a generalizable human electrophysiological signature of working memory load. *Psychophysiology*, *57*(12), e13691. https://doi.org/10.1111/psyp.13691

Awh, E., Barton, B., & Vogel, E. K. (2007). *Visual Working Memory Represents a Fixed Number of Items Regardless of Complexity*. *18*(7), 7.

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., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 780–790. https://doi.org/10.1037/0096-1523.24.3.780

Awh, E., Vogel, E. K., & Oh, S.-H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*(1), 201–208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Balaban, H., Drew, T., & Luria, R. (2019). Neural evidence for an object-based pointer system underlying working memory. *Cortex*, *119*, 362–372. https://doi.org/10.1016/j.cortex.2019.05.008

Bays, P. M. (2018). Reassessing the Evidence for Capacity Limits in Neural Signals Related to Working Memory. *Cerebral Cortex*, *28*(4), 1432–1438. https://doi.org/10.1093/cercor/bhx351

Bouchacourt, F., & Buschman, T. J. (2019). A Flexible Model of Working Memory. *Neuron*, *103*(1), 147-160.e8. https://doi.org/10.1016/j.neuron.2019.04.020

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and Reconstructing Color from Responses in Human Visual Cortex. *Journal of Neuroscience*, *29*(44), 13992–14003. https://doi.org/10.1523/JNEUROSCI.3577-09.2009

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional Templates in Visual Working Memory. *Journal of Neuroscience*, *31*(25), 9315–9322. https://doi.org/10.1523/JNEUROSCI.1097-11.2011

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 109. https://doi.org/10.1037/0096-1523.21.1.109

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142. https://doi.org/10.1146/annurev-psych-010814-015031

Diaz, G. K., Vogel, E. K., & Awh, E. (2021). Perceptual Grouping Reveals Distinct Roles for Sustained Slow Wave Activity and Alpha Oscillations in Working Memory. *Journal of Cognitive Neuroscience*, *33*(7), 1354–1364. https://doi.org/10.1162/jocn_a_01719

Drew, T., & Vogel, E. K. (2008). Neural Measures of Individual Differences in Selecting and Tracking Multiple Moving Objects. *Journal of Neuroscience*, *28*(16), 4183–4191. https://doi.org/10.1523/JNEUROSCI.0556-08.2008

Duncan, J. (1980). *The Locus of Interference in the Perception of Simultaneous Stimuli*.

Feldmann-Wüstefeld, T. (2021). Neural measures of working memory in a bilateral change detection task. *Psychophysiology*, *58*(1), e13683. https://doi.org/10.1111/psyp.13683

Felleman, D. J., & Van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex*, *1*(1), 1–47. https://doi.org/10.1093/cercor/1.1.1-a

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

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology*, *115*(1), 168–177. https://doi.org/10.1152/jn.00860.2015

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2017). Alpha-Band Oscillations Enable Spatially and Temporally Resolved Tracking of Covert Spatial Attention. *Psychological Science*, *28*(7), 929–941. https://doi.org/10.1177/0956797617699167

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. https://doi.org/10.3758/17.5.673

Fukuda, K., & Woodman, G. F. (2015). Predicting and Improving Recognition Memory Using Multiple Electrophysiological Signals in Real Time. *Psychological Science*, *26*(7), 1026–1037. https://doi.org/10.1177/0956797615578122

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349. https://doi.org/10.1152/jn.1989.61.2.331

Fuster, J. M., & Jervey, J. P. (1981). Inferotemporal Neurons Distinguish and Retain Behaviorally Relevant Features of Visual Stimuli. *Science*, *212*(4497), 952–955. https://doi.org/10.1126/science.7233192

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, *14*(3), 477–485. https://doi.org/10.1016/0896-6273(95)90304-6

Günseli, E., Fahrenfort, J. J., van Moorselaar, D., Daoultzis, K. C., Meeter, M., & Olivers, C. N. L. (2019). EEG dynamics reveal a dissociation between storage and selective attention within working memory. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-49577-0

Gunseli, E., Meeter, M., & Olivers, C. N. L. (2014). Is a search template an ordinary working memory? Comparing electrophysiological markers of working memory maintenance for visual search and recognition. *Neuropsychologia*, *60*, 29–38. https://doi.org/10.1016/j.neuropsychologia.2014.05.012

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*.

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2021). Controlling the Flow of Distracting Information in Working Memory. *Cerebral Cortex (New York, NY)*, *31*(7), 3323–3337. https://doi.org/10.1093/cercor/bhab013

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), Article 7238. https://doi.org/10.1038/nature07832

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral Delay Activity Provides a Neural Measure of the Number of Representations in Visual Working Memory. *Journal of Neurophysiology*, *103*(4), 1963–1968. https://doi.org/10.1152/jn.00978.2009

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. https://doi.org/10.1016/0010-0285(92)90007-O

Luck, S. J., & Vogel, E. K. (1997). *The capacity of visual working memory for features and conjunctions*. *390*, 3.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*(6601), 616–618. https://doi.org/10.1038/383616a0

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100–108. https://doi.org/10.1016/j.neubiorev.2016.01.003

McNab, F., Leroux, G., Strand, F., Thorell, L., Bergman, S., & Klingberg, T. (2008). Common and unique components of inhibition and working memory: An fMRI, within-subjects investigation. *Neuropsychologia*, *46*(11), 2668–2682. https://doi.org/10.1016/j.neuropsychologia.2008.04.023

Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review. *Journal of Cognition*, *2*(1), 36. https://doi.org/10.5334/joc.58

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the “strong-object” hypothesis. *Perception & Psychophysics*, *64*(7), 1055–1067. https://doi.org/10.3758/BF03194756

Poo, M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., Rudenko, A., Tsai, L.-H., Tsien, R. W., Fishell, G., Mullins, C., Gonçalves, J. T., Shtrahman, M., Johnston, S. T., Gage, F. H., Dan, Y., Long, J., Buzsáki, G., & Stevens, C. (2016). What is memory? The present state of the engram. *BMC Biology*, *14*, 40. https://doi.org/10.1186/s12915-016-0261-6

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*(1), 65–97. https://doi.org/10.1016/0010-0277(89)90014-0

Pylyshyn, Z. W. (2009). Perception, representation, and the world: The FINST that binds. In *Computation, cognition, and Pylyshyn* (pp. 3–48). MIT Press.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197. https://doi.org/10.1163/156856888X00122

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. *Nature Neuroscience*, *22*(8), Article 8. https://doi.org/10.1038/s41593-019-0428-x

Rihs, T. A., Michel, C. M., & Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by α-band EEG synchronization. *European Journal of Neuroscience*, *25*(2), 603–610. https://doi.org/10.1111/j.1460-9568.2007.05278.x

Samaha, J., Sprague, T. C., & Postle, B. R. (2016). Decoding and Reconstructing the Focus of Spatial Attention from the Topography of Alpha-band Oscillations. *Journal of Cognitive Neuroscience*, *28*(8), 1090–1097. https://doi.org/10.1162/jocn_a_00955

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. https://doi.org/10.1111/j.1467-9280.2009.02276.x

Shiffrin, R. M., & Schneider, W. (1977). *Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory*. 64.

Sprague, T. C., & Serences, J. T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*, *16*(12), 1879–1887. https://doi.org/10.1038/nn.3574

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. https://doi.org/10.3758/s13414-014-0633-3

Thyer, W., Adam, K. C. S., Diaz, G. K., Velázquez Sánchez, I. N., Vogel, E. K., & Awh, E. (2022). Storage in Visual Working Memory Recruits a Content-Independent Pointer System. *Psychological Science*, *33*(10), 1680–1694. https://doi.org/10.1177/09567976221090923

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*(6984), 751–754. https://doi.org/10.1038/nature02466

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural Limits to Representing Objects Still within View. *Journal of Neuroscience*, *33*(19), 8257–8263. https://doi.org/10.1523/JNEUROSCI.5348-12.2013

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. https://doi.org/10.1016/j.cogpsych.2014.01.003

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2015). Working Memory Delay Activity Predicts Individual Differences in Cognitive Abilities. *Journal of Cognitive Neuroscience*, *27*(5), 853–865. https://doi.org/10.1162/jocn_a_00765

Vaina, L. M. (1994). Functional Segregation of Color and Motion Processing in the Human Visual Cortex: Clinical Evidence. *Cerebral Cortex*, *4*(5), 555–572. https://doi.org/10.1093/cercor/4.5.555

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1656. https://doi.org/10.1037/0096-1523.24.6.1656

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., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. https://doi.org/10.1038/nature04171

Williams, M., Pouget, P., Boucher, L., & Woodman, G. F. (2013). Visual–spatial attention aids the maintenance of object representations in visual working

memory. *Memory & Cognition*, *41*(5), 698–715. https://doi.org/10.3758/s13421-013-0296-7

Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*(7080), 91–95. https://doi.org/10.1038/nature04262

Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, *13*(4), 167–174. https://doi.org/10.1016/j.tics.2009.01.008

Zeki, S. M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, *274*(5670), Article 5670. https://doi.org/10.1038/274423a0