

UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

Real-World Objects Have Greater Spatial Precision than Simple Stimuli in Visual Working Memory

Permalink

<https://escholarship.org/uc/item/1jn9v5t7>

Author

Lam, Kelvin

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA BARBARA

Real-World Objects Have Greater Spatial Precision than Simple Stimuli in Visual Working
Memory

A Thesis submitted in partial satisfaction of the requirements for the degree
Master of Arts in Psychological & Brain Sciences

by

Kelvin Lam

Committee in charge:

Professor Thomas C. Sprague, Chair

Professor Mary Hegarty

Professor Barry Giesbrecht

September 2022

The thesis of Kelvin Lam is approved.

Mary Hegarty

Barry Giesbrecht

Thomas C. Sprague, Chair

August 2022

BSTRACT

Real-World Objects Have Greater Spatial Precision than Simple Stimuli in Visual Working

Memory

by

Kelvin Lam

Much of what we know about visual working memory has been uncovered using simple stimuli such as colored squares, but recent studies have suggested that alternative conclusions can be drawn using more complex stimuli, such as real-world objects. A possible explanation as to why these differences exist is that objects are redundantly coded broadly across several brain regions, whereas simple stimuli like colored squares are encoded in a more limited set of regions. We hypothesize that spatial position is enhanced automatically by this redundant encoding. In our experiment, participants (N=30) were given a delayed spatial recall task in which participants remembered arrays of either real world objects or colored squares. Participants maintained the precise spatial position of 1, 2, or 6 visual stimuli over a brief 1.5s delay period. We found that participants recalled spatial positions of objects with higher precision than the colored squares. Moreover, in separate trials where participants discriminated which of two stimuli appeared at a probed location, performance was identical between stimulus conditions. Altogether, this result supports the notion that incidental features of real-world objects, such as their spatial position, can be remembered with greater precision than those of simplistic stimuli typically used in laboratory tests.

Introduction

Working memory (WM) is a core component of human cognition that is typically conceptualized as buffering information before encoding into long term memory (Baddeley & Hitch, 1974), but it also functions as our ability to maintain information to do certain tasks over a short period of time. Visual WM is a limited resource system which decreases with information load (Luck & Vogel, 1997), has a maximum capacity (Miller, 1956; Luck & Vogel, 1997; Vogel et al., 2001; Cowan, 2001) that varies based on the individual (Engel et al., 1999; Vogel & Machizawa, 2005). Originally conceptualized as 7 items (Miller, 1956), modern estimates place this capacity at around 4 discrete items once tested using stimuli that cannot be chunked (Vogel et al., 2001; Cowan, 2001). This estimate was further supported by neuroimaging data in EEG and fMRI (Vogel & Machizawa, 2004; Todd & Marois, 2004).

Despite numerous papers reiterating capacity estimates converge at around 4 items (Luck & Vogel, 1997; Vogel et al., 2001; Vogel & Machizawa, 2004; Awh et al., 2007; Adam et al., 2017; Quirk et al., 2020), it is also possible that WM capacity is instead dependent on both the stimuli and testing modality used to investigate these properties. First, there is evidence to suggest that when the remembered stimuli are more visually complex (such as shaded cubes), thus more cognitively taxing to maintain in WM, capacity estimates fall as low as 2 items (Alvarez & Cavanaugh, 2004). Additionally, there is evidence presenting higher capacity estimates when stimuli are naturalistic real-world objects, possibly because semantic information acts as an organizing device to reduce cognitive load (Asp et al., 2021; Brady et al., 2016; Brady et al., 2020).

In addition to possibly capacity differences, simple stimuli and naturalistic stimuli also differ in the way they are represented on the cortex. Viewing colors activates early visual cortex, specifically areas V1-3, hV4 (Brewer et al., 2005, Brouwer & Heeger, 2009). In comparison, viewing objects activates visual cortex, areas in the parietal lobe, and some ventral temporal areas as well, all of which having some degree of spatial selectivity (Haxby et al., 2001; Serences, 2016; Grill-Spector et al., 2001). If it is the case that WM recruits cortical regions implicated during perception (Harrison & Tong, 2009; Serences et al., 2009) distributed and redundant coding may play a part in explaining previously discussed WM benefits (Stojanoski et al., 2019). What benefit might these redundant representations provide? One feature of visual cortex is that it has a very precise retinotopically organization which underlies our spatial perception (Hubel & Wiesel, 1962; Engel et al., 1997). Even though the parietal and ventral temporal areas have a lower spatial resolution than the visual cortex, it might be the case that just having that redundant spatial encoding improves the spatial precision as compared to colored squares. Although redundant representations arise necessarily from the architecture of neural connections and may not indicate a mechanism for an unknown function, there is evidence that points to some serendipitous and unexpected benefits to their existence. A recent study has shown that the existence of multiple visual representations in WM can help protect from information degradation due to distractors (Hallenbeck et al., 2021).

Given that there is reason to believe that naturalistic and simple stimuli behave differently in WM and that these stimuli types have quantifiably different neural activation patterns, these experiments were used to determine if any other behavioral differences

existed, namely in the domain of spatial working memory. First, there is evidence that suggests that semantic information can organize perceptual information and reduce cognitive load (e.g., Asp et al., 2021), and also a body of work that shows that precision is inversely related to cognitive load. However, it is not entirely clear if spatial location is one of the features of a visual stimulus which can be organized more efficiently with semantic information as a guide. Another reason why objects may show better spatial precision may also lie with the spatiotopic organization of visual areas and spatial precision could be incidentally enhanced due to redundant encoding. For these reasons, we hypothesize that the spatial position of real-world objects will be more precisely reported as compared to the positions of simple stimuli.

Methods

Participants

The final sample contained 30 participants (18 male & 12 female; mean age = 19.8) from University of California, Santa Barbara and were compensated \$10 upon successful completion of the experiment. 1 participant was excluded for repeatedly interrupting the experiment to try and start a conversation with the experimenter. Participants provided informed consent and all procedures complied with guidelines established by University of California, Santa Barbara institutional review board.

Stimulus & Procedure

We designed the task using the Brady et al. 2016 task in mind to be able to attempt a conceptual replication. Participants were seated 52 centimeters from the screen, head fixed.

Stimuli were presented on a 54 cm x 32 cm Dell monitor, at 2560x1440 resolution and refresh rate at 60 Hz. The stimulus presentation script was written and executed using MATLAB Psychtoolbox.

Participants completed 8 blocks of 30 runs, for a total of 240 trials. Participants were instructed to view a memory array consisting of either colored squares or real-world objects, to perform on of two separate tasks (*fig. 1*). These tasks were counterbalanced by experiment half— participants were randomly assigned to do the 2AFC task blocks first, or the continuous report task blocks first. For both tasks, the stimulus display was presented the same way. On every trial, participants were shown a circle 16 degrees visual angle (DVA) in diameter, centered on the screen. A fixation cross was presented in the middle of the screen, and participants were instructed to keep their gaze fixated on the cross for the entire block. After 500 ms, the trial begins, and participants are presented with the stimuli. Each stimulus display was either all colored squares or all real-world objects, with a set size of 1, 2, or 6.

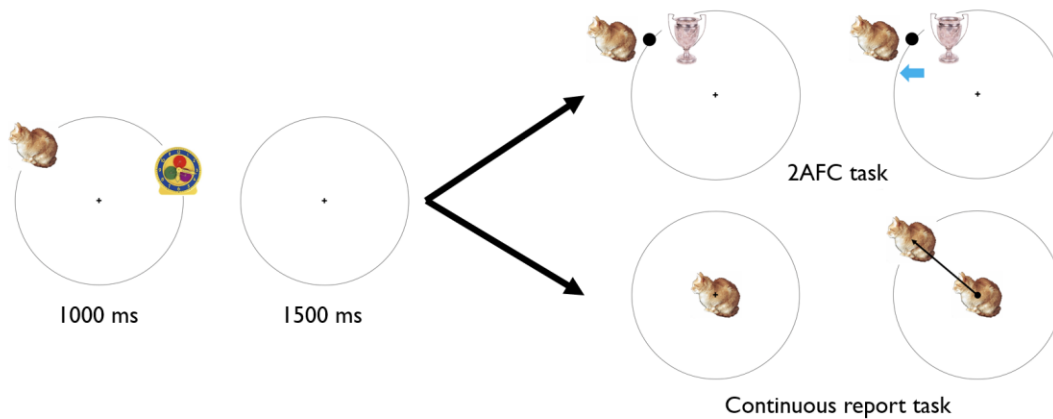


Figure 1. Task. Participants are shown an array of items on a centrally presented ring, and asked to remember the display. The test was either a 2AFC task, where participants were probed with a location and asked to identify which of 2 items they recall being presented at that location. During the continuous report task, participants are shown one of the memory items, and asked to use the mouse to drag the item to the remembered location as accurately as possible.

The stimuli were presented in different locations on the circle for 1000 ms, due to evidence that a significant difference between capacity estimates of the object and colored squares emerges after 1000 ms (Brady et al., 2016).

The set size of images presented varied by trial. On 40% of trials, participants are asked to remember 1 image, 40% of trials 2 images, and 20% of trials 6 images. On half the trials shown, participants were presented with objects (courtesy of Brady et al., 2008), and on the other half of trials the stimuli were colored squares. On object trials which had a set size of 2 or 6, the objects were selected from unique, non-overlapping categories. We intended the different image in the set to be categorically different so that we do not see performance being influenced due to image similarity. Likewise, colors selected using the HSV colorspace were at least 20 degrees apart. Both types of stimuli were 2.4 DVA in diameter and presented on pseudo-randomly generated locations on a circle indicating the decision space. The positions were generated by picking 6 equidistant locations on the circular response space, adding a random amount of jitter to each location, and rotating all locations by a random amount. On trials where the set size was less than 6, each stimulus to be presented at one of these generated locations at random. We wanted to focus the set size distribution to lower set sizes, because if it is true that a difference in processing efficacy leads to performance differences, above capacity trials in both stimulus conditions would hit a processing ceiling and not show an effect (Luck & Vogel, 1997). Moreover, set size 6 was primarily included to match the methods in Brady et al., 2016.

After the display, the stimuli were removed from the display and participants were shown a blank screen for 1500 ms. Participants would either complete the continuous report task or the 2AFC task. On continuous report trials, participants were given the identity of one of the images on the screen and were prompted to identify the location where they last remembered seeing that image. They indicated their choice by dragging the image of the stimulus to the correct location and clicking to indicate their response. Participant accuracy on each continuous report task trial was recorded as the angular difference between reported location and the actual location that stimulus was presented. Participants had a 3 second window to make a response. At the end of a continuous report block, the participants were given feedback on their behavioral performance in the form of the mean angular difference over all trials in the block.

On the 2AFC trials, participants were presented with one of the locations on the circle, indicated by a circle appearing at the intended location. Two images were presented to the left and right of the circle, respectively. One item was the correct memory item, and the other was a foil which was categorically different than any item shown in the display. Similarly, for the colored square condition, the color was different than any of the colors seen in the display. The participant was asked to pick which one of these images was shown at the probed location, and indicated their choice using the left/right arrow keys, corresponding to the left or right image. At the end of a 2AFC block, participants were given feedback in the form of a percentage of correctly answered trials in the block.

Behavioral Analysis

The results of the continuous report task were analyzed using standard mixture model (Zhang & Luck, 2008; Suchow et al. 2013). The standard mixture model parameterizes the distribution of response deviations, measured in degrees polar angle around the circular response space, and outputs the data's response precision, and the guess rate (*fig. 2*).

The model assumes that there are two distributions which contribute to a response histogram, a distribution when the item is in memory and a distribution for when subjects are guessing. The guess distribution is conceptualized as participants having no memory of the images shown in the display, so instead of an accurate response, the participants would click a location on the response space at random. Over the span of many trials, these random responses will end up resembling a uniform distribution, which we can subtract from the original data. The proportion of responses that constitute this uniform distribution is considered the guess rate. After subtracting the uniform distribution out of the data, we are left with what appears to be a circular normal (von Mises) distribution. Precision is measured as the standard deviation of that distribution. Precision is inversely proportional to the standard deviation (i.e., when the standard deviation decreases, responses are becoming more precise). For the 2AFC data, we are going to be analyzing the proportion of trials the participant responded correctly, which we will refer to just as accuracy, and their reaction time on these trials, measured from the period where the test stimuli are presented until the button press which indicated their decision.

For the modelling, set size 6 trials were removed due to having too few observations for the mixture model to converge. Because the mixture model attempts to fit a circular normal distribution to the errors, if there are too few trials the model will be poorly fit and return inaccurately calculated parameters.

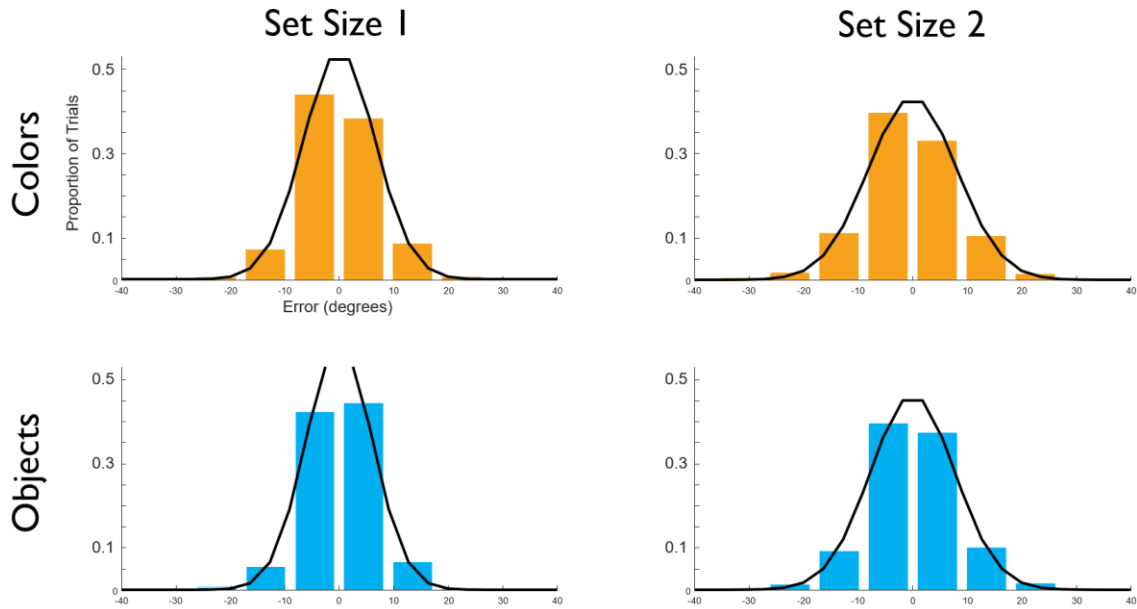
Each participant received 4 precision and 4 guess scores, one for each of the 4 stimulus conditions. A 2 (stimulus type; objects & colored squares) x 2 (set size; 1 & 2) ANOVA was performed on both the precision scores and the guess scores. For the 2AFC task, we did not remove set size 6 due to the analysis being a simple accuracy calculation. Each participant had 6 accuracy scores, one corresponding to each condition. A 2 (stimulus type; objects & colored squares) x 3 (set size; 1, 2, 6) ANOVA was performed on the 2AFC accuracy and reaction time.

Results

Continuous Report Task

Running a 2 way ANOVA on set size and stimulus type, we see a main effect of set size: the standard deviation of responses on set size 2 trials was higher than the standard deviation of the set size 1 trials (2 way ANOVA, $F(1,119)=40.4, p<.001$) (*fig. 3a*). In other words, set size 2 was on average, less precise than set size 1 trials. We also found a main effect of stimulus type: images of objects had a smaller standard deviation as compared to the colored squares, suggesting that the spatial memory of objects is more precise than the spatial memory of colored squares (2 way ANOVA, $F(1,119)=7.55, p=.027$). The interaction term for these effects was non-significant (2 way ANOVA, $F(1,119)=1.32, p=.25$). For the guess

Figure 2. Response error histograms, across subjects. Raw data from each of the 4 conditions, combined across subjects. The standard mixture model aggregates the response errors across a given condition for a given participant, and reports the standard deviation of the fitted von Mises distribution as a measure of precision.



rate, we do not find any effect (2 way ANOVA, stimulus type: $F(1,119)=0.44, p=.51$; set size: $F(1,119)=2.57, p=.12$) (fig. 2b). We might be seeing an emerging difference in guess rate between objects and colored squares as set size increases, but this difference at set size 2 is due to 1 extreme outlier.

2AFC Task

In the 2AFC task, we see a main effect of set size on participant accuracy. There was a main effect of set size (2 way ANOVA, $F(2,179)=238.3, p<.001$), but not of stimulus type (2 way ANOVA, $F(1,179)=3.64, p=.066$) (fig. 3c). Post hoc comparisons revealed that set size 6 was significantly lower in accuracy than set size 1 and 2, and set size 1 and 2 not being significantly different from one another. For the reaction time data, we see a complementary

set of data. We see that there is a main effect of set size on reaction time (2 way ANOVA, $F(1,179)=51.82, p<.001$), and no effect of stimulus type (2 way ANOVA, $F(1,179)=0.42, p=.52$). Post hoc tests revealed that set size 6 trials were significantly slower than set size 1 and 2 trials, set size 1 and 2 not being significantly different from one another.

Discussion

This study investigated the differences in spatial precision between real world objects and simple stimuli with a continuous spatial report task and attempted to replicate a finding from Brady et al. (2016). This experiment confirmed the hypothesis that doing a spatial report task with real world objects resulted in higher precision results as compared to simple stimuli. However, in the second finding of that study, we were unable to replicate the results from Brady et al., (2016). The finding that objects are remembered more spatially precise in working memory is novel in and of itself.

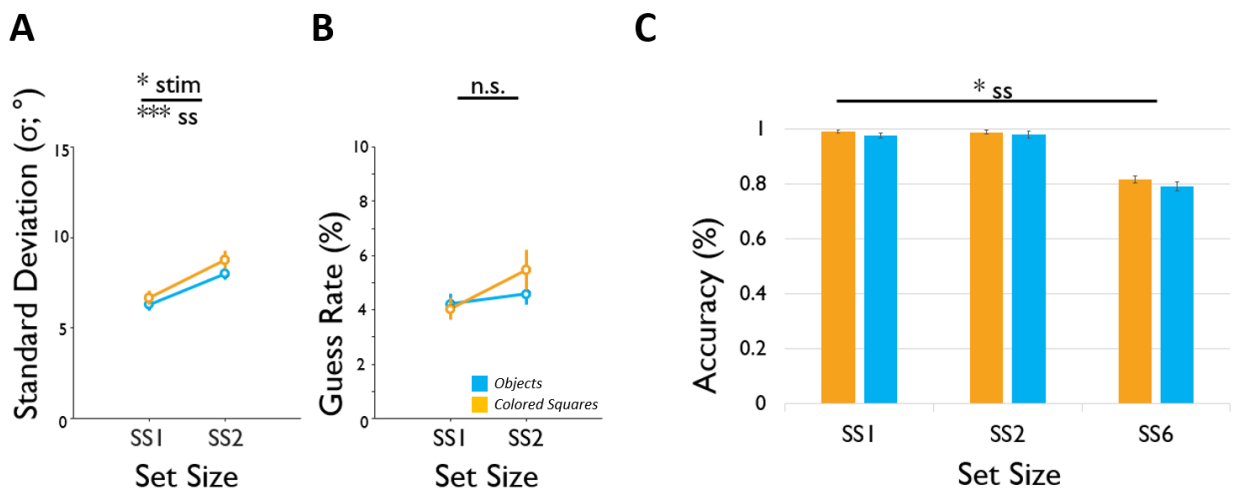


Figure 3. Real-world objects are more spatially precise than colored squares in working memory. A) The standard deviation (the inverse of precision) of colored squares and real-world objects across set size during the continuous report task. B) Percentage of trials where participants were determined to be guessing during the continuous report task. C) Proportion of correct trials during the 2AFC task.

Unseen Effects of Semantic Information

Integral to understanding the debate about capacity estimate are the slot model of working memory and the resource model of working memory. Slot-based models assume that there are a discrete number of slots in which you can hold things in working memory, no matter the complexity. This is backed by papers showing participants have a fixed capacity estimate under different stimulus types, with apparently corroborating neurophysiology (e.g. Luck & Vogel, 1997; Vogel & Machizawa, 2004). Resource models assume that there is a fixed amount of cognitive resource that can be allocated to items in working memory, without an individual maximum that it can allocated to. Instead, WM is limited by how much resource you can allocate to each item, with decreasing fidelity as you give each item less resource. This is backed by papers showing that capacity or memory fidelity is affected by the types of stimuli being stored in memory (e.g. Alvarez & Cavanaugh, 2007; Brady et al., 2016).

In this experiment, we find that the spatial memory for real world objects is more precise than for colored squares. This could suggest that real world objects receive more cognitive resource, and therefore have a more precise representation. Additionally, articles that argue in favor of real-world objects having special properties are those also arguing in favor of resource-based models. Although not having been observed before, it is clear to see how extra cognitive processing would support higher fidelity spatial representations. For instance, when working memory items have less access to resources (such as by increasing the number of remembered items), memory fidelity decreases (Zhang & Luck, 2008).

However, it is not exactly clear why real-world objects would be granted extra cognitive resources if in this task, they are equally as behaviorally relevant as the colored squares. Instead of extra resources, perhaps this should be understood as a function of reduced load instead of further processing. Considering there is evidence to believe that access to semantic information can help organize perceptual information to reduce load (Asp et al., 2021), an unexpected consequence of that reorganization could be this spatial precision benefit we observe. By reducing the load, it could be the case that the same quantity of cognitive resources is allocated regardless of stimulus type, but the real-world objects are more able to efficiently use those resources and resulting in higher memory fidelity.

Failure to replicate

The 2AFC task in this experiment is a conceptual replication of the task in Brady et al., (2016). As reported, we were unable to find the increased accuracy for real world objects reported in their study. This is not a unique issue (Quirk et al., 2020), and the original authors have already put out a rebuttal, which claims the which stimuli chosen as the foil during the 2AFC task has an impact on finding a significant effect (Brady & Stoermer, 2020). Despite not replicating their results, our findings are consistent with their proposed explanation that real-world objects have decreased load as compared to colored squares.

Future Directions

One possible direction forward is to investigate the role of redundant visuo-spatial representations across the cortex in determining this real-world object benefit. To accomplish this, one analysis to consider is to consider if real world objects also show a correlation

between their spatial precision and decoding strength (as seen in Emrich et al., 2013).

Although this result would be interesting, it has been observed that models trained on more areas tend to have better decoding strength (e.g., Harrison & Tong, 2009); it is too early to speculate if something noteworthy or just a function of how encoding models are constructed.

References

- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106-111.
- Asp, I. E., Störmer, V. S., & Brady, T. F. (2021). Greater visual working memory capacity for visually matched stimuli when they are perceived as meaningful. *Journal of Cognitive Neuroscience, 33*(5), 902-918.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science, 18*(7), 622-628.
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision, 11*(5), 4-4.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325-14329.
- Brady, T.F. and Störmer, V.S. (in press). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*.
- Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences, 113*(27), 7459-7464
- Brewer, A. A., Liu, J., Wade, A. R., & Wandell, B. A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience, 8*(8), 1102-1109.
- Brouwer, G. J., & Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience, 29*(44), 13992-14003.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-114.

- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, 7(2), 181-192.
- 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.
- Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, 20(2), 177-182.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10-11), 1409-1422.
- Hallenbeck, G. E., Sprague, T. C., Rahmati, M., Sreenivasan, K. K., & Curtis, C. E. (2021). Working memory representations in visual cortex mediate distraction effects. *Nature Communications*, 12(1), 1-18.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632-635.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425-2430.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106-154.
- Kerr, J. L. (1971). Visual resolution in the periphery. *Perception & Psychophysics*, 9(3), 375-378.
- Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General*, 145(1), 95.
- Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision*, 17(6), 20-20.

- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279-281.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25
- Schneegans, S., & Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. *Journal of Neuroscience*, 37(14), 3913-3925.
- Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, 128, 53-67.
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207-214.
- Sprague, T. C., Ester, E. F., & Serences, J. T. (2014). Reconstructions of information in visual spatial working memory degrade with memory load. *Current Biology*, 24(18), 2174-2180.
- Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of vision*, 14(12), 6-6.
- Stojanoski, B., Emrich, S. M., & Cusack, R. (2019). Representation of semantic information in ventral areas during encoding is associated with improved visual short-term memory. *bioRxiv*.
- Suchow, J. W., Brady, T. F., Fournie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, 13(10):9, 1–8. doi:10.1167/13.10.9.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6(2), 171-178.
- Quirk, C., Adam, K. C. S., & Vogel, E. K. (2020). No evidence for an object working memory capacity benefit with extended viewing time. *Eneuro*, 7(5)

- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984), 748-751.
- Vogel E.K., Woodman G.F., Luck S.J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 92–114.
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39(2), 293-306.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233-235.