

Limits on Perceptual Encoding Can Be Predicted From Known Receptive Field Properties of Human Visual Cortex

Michael A. Cohen

Massachusetts Institute of Technology and Harvard University

Juliana Y. Rhee and George A. Alvarez

Harvard University

Human cognition has a limited capacity that is often attributed to the brain having finite cognitive resources, but the nature of these resources is usually not specified. Here, we show evidence that perceptual interference between items can be predicted by known receptive field properties of the visual cortex, suggesting that competition within representational maps is an important source of the capacity limitations of visual processing. Across the visual hierarchy, receptive fields get larger and represent more complex, high-level features. Thus, when presented simultaneously, high-level items (e.g., faces) will often land within the same receptive fields, while low-level items (e.g., color patches) will often not. Using a perceptual task, we found long-range interference between high-level items, but only short-range interference for low-level items, with both types of interference being weaker across hemifields. Finally, we show that long-range interference between items appears to occur primarily during perceptual encoding and not during working memory maintenance. These results are naturally explained by the distribution of receptive fields and establish a link between perceptual capacity limits and the underlying neural architecture.

Keywords: visual cognition, visual working memory, receptive fields, capacity limits, attention

The processing limitations of human cognition are well established: across several modalities (e.g., vision, audition, and somatosensation), our ability to perceive, attend to, and remember multiple items is surprisingly limited (Cowan, 2001; Gallace et al., 2006; Vitevitch, 2003). Over several decades, many cognitive models have been developed to explain these limitations (Baddeley, 1998; Broadbent, 1958; Deutsch & Deutsch, 1963; Kahneman, 1973). In the domain of visual cognition, contemporary models of attention and working memory have characterized these capacity limitations as either a fixed limit on the number of discrete representations (Buschman, Seigel, Roy, & Miller, 2011; Cowan, 2001; Zhang & Luck, 2008), or as a finite, flexible pool of resources (Alvarez & Franconeri, 2007; Bays & Husain, 2008; Holcombe &

Chen, 2012). Here, we take a different approach to understanding these limitations by focusing on the functional organization of the ventral visual stream into representational maps (Franconeri, Alvarez, & Cavanagh, 2013). More specifically, we examine how the size, distribution, and response properties of receptive fields constrain perception.

A given sensory neuron will only respond to stimuli that activate certain receptors, which together comprise that neuron's receptive field. For example, a neuron in area V1 only responds to light that falls on a particular set of photoreceptors in the retina. Receptive fields exist in vision (Wandell et al., 2007), audition (Knudsen, 1982), somatosensation (Graziano & Gross, 1993), and possibly olfaction (Wilson, 2001). Without question, receptive fields have been most extensively studied in the ventral pathway of the primate visual system. Decades of research have revealed several properties of visual receptive fields. First, receptive fields get larger at each successive stage of processing. While receptive fields in V1–V4 range from $\sim 2^\circ$ to $\sim 6^\circ$ in the periphery (Gattass et al., 1988; Winawer et al., 2010), peripheral inferotemporal (IT) receptive fields are on average $\sim 10^\circ$ and can be up to $\sim 26^\circ$ (Distler et al., 1993; Op de Beeck & Vogels, 2000). Second, both early visual and IT receptive fields have a strong contralateral bias, though receptive fields beyond V4 do often cross the vertical meridian (Golomb & Kanwisher, 2012; Kravitz et al., 2008; Wandell et al., 2007). Finally, when multiple objects land in the same receptive fields, those items compete for the neuron's response (Beck & Kastner, 2005; Desimone & Duncan, 1995). Unless attention is deployed, the neuron's output will be the average of the response to the individual items had they been presented individually (Zoccolan, Cox, & DiCarlo, 2005).

From this, we predict that when multiple items are close enough in space to land in the same receptive fields, competition between

This article was published Online First August 31, 2015.

Michael A. Cohen, McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, and Department of Psychology, Harvard University; Juliana Y. Rhee, Department of Molecular and Cellular Biology, Harvard University; George A. Alvarez, Department of Psychology, Harvard University.

Thanks to Arash Afraz, Nancy Kanwisher, Ken Nakayama, Maryam Vaziri-Pashkam, and Jeremy Wolfe for comments on the project, and to Tim Brady, Daryl Fougny, and Jordan Suchow for comments on an earlier version of the manuscript, and Sarah Cormeia for help with data collection. This work was supported by an NSF-GRFP and NIH-NRSA (F32EY024483) to Michael A. Cohen and NSF CAREER (BCS-0953730) to George A. Alvarez.

Correspondence concerning this article should be addressed to Michael A. Cohen, McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, 46-4141, Cambridge, MA 02139. E-mail: michaelthecohen@gmail.com

those items will lead to degraded perceptual representations and decreased performance on a behavioral task. If those same items are repositioned to make them less likely to land within the same receptive fields, behavioral performance will improve. Critically, since receptive fields vary in size and in what they respond to, we predict that these spatial interference effects will vary depending on the items being processed and the receptive fields that represent the task-relevant features of those items. For instance, receptive fields that represent more high-level features (e.g., those found in faces, scenes, and objects) are much larger than receptive fields that represent low-level features (e.g., colors, orientations; Kravitz et al., 2008). Thus, we predict there will be long-range interference between items comprised of high-level features (e.g., faces), while only short-range interference will be found with low-level features (e.g., color).

Conceptually, the framework described here is broadly consistent with the reverse hierarchy theory of perceptual learning (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002). According to this theory, learning starts in higher-level visual regions and works backward until the region with the highest signal-to-noise ratio is identified. Here, we suggest that the neural regions with the most task-relevant perceptual representations (i.e., highest signal) will depend on the stimuli being processed. For example, we assume that with a task that requires encoding and comparing properties of higher-level objects (e.g., face identity), the most task-relevant representations will be in IT. Conversely, when performing the same task with lower-level features, the critical neural regions will no longer be in IT cortex, but will instead be in earlier visual cortex. Here, we suggest that because different parts of the visual system have differently sized receptive fields, the spatial interference patterns between items will vary as a function of the particular stimuli being processed.

To test this hypothesis, participants performed a change detection task with different types of items (i.e., high-level vs. low-level) placed in different spatial locations (spaced near vs. far, within vs. across hemifields). Overall, we found strong evidence that spatial interference between items depends on the items being presented, their distance from one another, and their relative locations in the visual field. These interference patterns were well predicted by the organization of receptive fields across the ventral stream, but are not predicted by several cognitive models of capacity limitations (Alvarez & Franconeri, 2007; Bays & Husain, 2008; Buschman et al., 2011; Cowan, 2001; Holcombe & Chen, 2012; Zhang & Luck, 2008). These results suggest that the organization of the ventral visual stream into spatial maps limits perception, specifically the ability to encode multiple items simultaneously (Franconeri et al., 2013).

Experiments 1a and 1b

Because high-level receptive fields are relatively large and are primarily confined to the contralateral hemifield, high-level items will likely compete with one another when presented within a visual hemifield (Figure 1a). When those items are presented across hemifields, but at the same relative spacing, there will be less competition within a given receptive field. Low-level receptive fields, meanwhile, are small enough that they will often only encompass one item at a time in these displays. Therefore, the amount of interference within a receptive field will be less affected

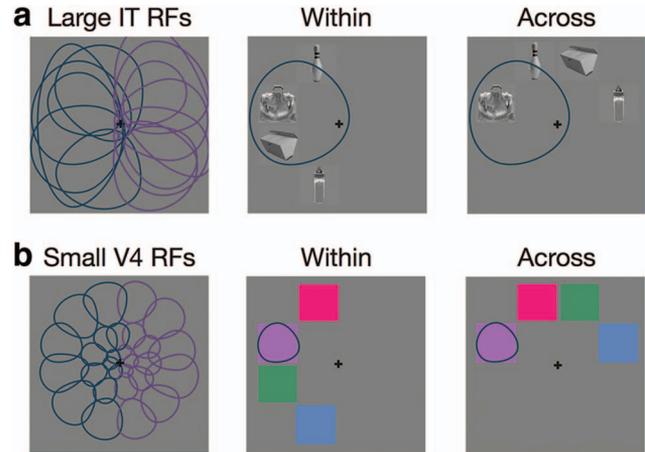


Figure 1. Idealized schematic depiction of receptive field size and distribution at different stages of the ventral visual pathway and the effects of items being placed in different spatial locations. (a) inferotemporal (IT) receptive fields (RF) are primarily confined to a visual hemifield and are large enough to encapsulate multiple items when they land within a hemifield. When those same items are presented across the hemifields, there are fewer instances of multiple items in one receptive field. (b) Receptive fields in lower-level cortical regions (e.g., V4) are also contralaterally confined but are often too small to encapsulate multiple items regardless of whether those items are presented within or across visual hemifield. To be more consistent with a variety of electrophysiology results, all receptive fields are drawn to have a slight cone-like shape that is biased toward or includes part of the fovea (Motter, 2009; Op de Beeck and Vogels, 2000).

by within-versus across-hemifield presentation (Figure 1b). Thus, we predicted that there would be hemifield effects for high-level, but not low-level, visual stimuli.

Method

Equipment. For all experiments, stimuli were presented on a 24-inch LCD monitor with a 60 Hz refresh rate. The position of the left eye was monitored with a video-based, desk-mounted ISCAN 18937 eye-tracker sampling at 60 Hz. A chin and forehead rest was used to minimize head movements and maintain a constant viewing distance of 57 cm. The experiments were created and controlled on a computer running MATLAB with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Eye position was streamed in real time from the eye-tracker to the computer running MATLAB. Experimental trials would begin only if participants were fixating on a centralized cross and immediately stopped if fixation was broken at any point during the trial. If participants broke fixation, the trial would restart, maintaining the assigned condition but with new, randomly selected stimuli. The background color of the screen was gray, with a luminance of 68.5 cd/m². All responses were made on a keyboard.

Participants. All participants were recruited from the Harvard University community and reported normal or corrected-to-normal vision. There were 12 participants for Experiment 1a and two separate groups of 12 participants used for Experiment 1b (one group performed the task with colors, another group with Gabors).

Stimuli. In Experiment 1a, stimuli were grayscale pictures of faces, scenes, and objects matched within category for luminance and contrast. Faces were photographs of Caucasian individuals looking directly at the camera, smiling, and facing forward. Scene stimuli were photographs of buildings, mountains, highways, and beaches. Both the face and scene stimuli came from a previously used stimulus set (Cohen et al., 2014). Objects were isolated on a gray background, matching the background of the screen, which had a luminance of 68.5 cd/m².

In Experiment 1b, stimuli were colored squares and oriented Gabors. There were nine color squares evenly distributed around a quarter of the circle in *CIE L*a*b* color space centered at $L = 54$, $a = 18$, $b = -8$, with a radius of 59. Only a portion of the color space was used to increase the difficulty and avoid ceiling performance. The Gabors were presented at 100% contrast, 1.5 cycles per degree of visual angle, on a gray background with a luminance of 68.5 cd/m².

Within a trial, all items were drawn from the same category (e.g., four faces, or four Gabors). All stimuli were square 6° × 6° images. The center of all stimuli presented was 11.4° of visual angle away from the center fixation point. The center-to-center distance of any two adjacent stimuli was 8.7°. The faces, scenes, and colors filled the entire square. The objects were scaled to be as large as possible while still fitting within this square. Gabors were scaled so that the edges of the Gabor would fade into the background smoothly while still fitting within the 6° × 6° square.

Procedure. Participants performed a change detection task in which they had to detect a change between two successively presented displays. On each trial, items were presented for 800 ms, followed by a 500 ms blank fixation display, and then by a second, probe display, which remained until participants responded using the keyboard (Figure 2a). Visual feedback was immediately given to let the participant know if that trial was answered correctly.

Stimuli were arranged either across hemifields or within a hemifield (Figure 2b). Participants fixated on a central cross and eye position was monitored with an eye-tracker to ensure that fixation was maintained. In Experiment 1a, each of the three stimulus categories appeared equally often in each of the four configurations (above, below, left, or right of fixation). For each configuration/stimulus pairing, changes occurred on half

the trials and at each location of each configuration equally often. On change trials, one image was replaced by a new image that was not on the first display. All trial types were randomized within one block. In Experiment 1b, each of the configurations occurred equally often, with changes occurring on half of the trials and in each particular location equally often. With the color displays, a change trial would entail one color being replaced by another random color from the stimulus set. With the Gabor displays, a change trial would entail rotating one Gabor 90°.

In Experiment 1a, there were 10 practice trials and 432 experimental trials. In Experiment 1b, there were 10 practice trials and 312 experimental trials. There were eye-tracker calibration sequences that occurred before the practice trials, the experimental block, and halfway through the experimental trials.

Results and Discussion

Across versus within-hemifield performance was compared for all stimulus categories. The size and location of the stimuli were chosen such that high-level receptive fields would encompass multiple items within a hemifield more than across hemifields (Distler et al., 1993; Kravitz et al., 2008; Op de Beeck & Vogels, 2000). In contrast, most low-level (i.e., V1–V4) receptive fields would be too small to regularly encompass multiple items (~2° to ~8° at 11.4° in the periphery; Gattass et al., 1988; Winawer et al., 2010). Thus, we predicted that high-level items, but not low-level items, would be affected by different hemifield arrangements.

In Experiment 1a, performance was better on across-hemifield displays than within-hemifield displays for all high-level categories (faces across $d' = 1.07$, $SEM = 0.14$; faces within $d' = 0.69$, $SEM = 0.15$; $t(11) = 3.02$, $p < .05$; scenes across $d' = 1.48$, $SEM = 0.16$; scenes within $d' = 1.02$, $SEM = 0.14$; $t(11) = 3.50$, $p < .01$; objects across $d' = 2.32$, $SEM = 0.11$; objects within $d' = 1.88$, $SEM = 0.08$; $t(11) = 3.18$, $p < .01$; Figure 3). Furthermore, a 3 × 2 analysis of variance (ANOVA) revealed no difference in the size of the hemifield effect between the categories, $F(2, 10) = .09$, $p = .92$.

While these results are consistent with the receptive field hypothesis, they may simply stem from the existence of independent attentional resources within each hemifield (Alvarez & Cavanagh,

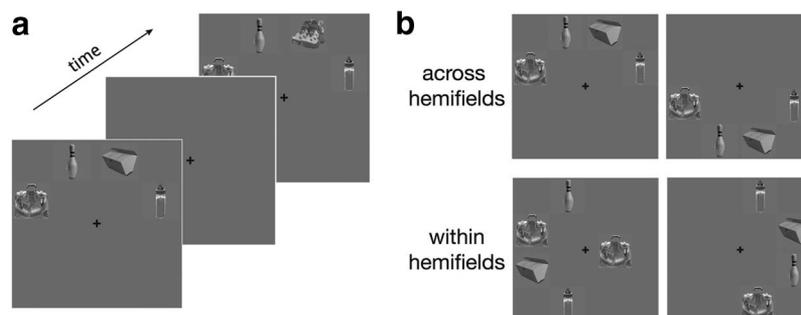


Figure 2. (a) Example of a single trial from Experiment 1a. In this case, objects are presented across hemifields, above the fixation point. (b) Stimuli could be arranged in one of four possible configurations: two across hemifields (above or below fixation) and two within a hemifield (left or right of fixation).

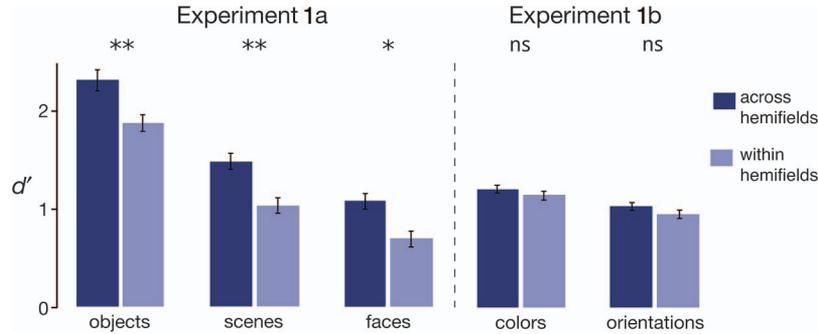


Figure 3. Results from Experiments 1a and 1b. Twelve participants performed Experiment 1a, another 12 participants performed Experiment 1b with colors, and another 12 performed Experiment 1b with orientations. Asterisks indicate significant differences ($p < .05$), within-subjects two-tailed t tests. Error bars reflect within-subject SEM (Cousineau, 2005).

2005; Clewenger & Beck, 2014; Delvenne, 2005). Experiment 1b served to test this possibility by replacing the high-level items with low-level items: colors and orientations. If the hemifield effect in Experiment 1a is because of separate attentional resource pools, this effect should also be observed with colors and orientations. However, if the hemifield effect is caused by interference within receptive fields, Experiment 1b should not show this effect because the displays were designed to minimize the relative number of instances in which multiple items land within the same V1–V4 receptive fields—areas that are highly sensitive to color and orientation (Figure 1b).

As predicted by the receptive field account, there was no hemifield effect for colors (across $d' = 1.19$, $SEM = 0.16$; within $d' = 1.14$, $SEM = 0.15$: $t(11) = 0.78$, $p = .45$) or orientations (across $d' = 1.02$, $SEM = 0.09$; within $d' = 0.94$, $SEM = 0.08$: $t(11) = 1.05$, $p = 0.32$) (Figure 3). To compare Experiments 1a and 1b, the results were averaged across high-level category for Experiment 1a (where each subject was tested in each condition, and there were no differences between conditions). We then determined the difference in performance for the within and between hemifield presentations. These difference scores were then compared separately to the difference in performance for the within and between hemifield presentations for both color and orientation groups of Experiment 1b. A between subjects t test in both cases showed a significant interaction between display configuration (across/within-hemifield) and stimulus type (high-/low-level) when the high-level items were compared with both colors, $t(11) = 4.52$, $p < .001$ and orientations, $t(11) = 4.18$, $p < .001$. These results demonstrate a clear asymmetry in the spatial interference patterns between high-level and low-level items that is predicted by differences in the size of receptive fields across the ventral stream.

Experiment 2

If the hemifield effect in Experiment 1a is the result of competition within large, high-level receptive fields, a similar effect should be found with low-level items if they are close enough to land within individual low-level receptive fields. While Experiment 1b and other previous attempts have failed to find a hemifield

effect with color (Delvenne, 2005; Mance et al., 2012), it is possible that items were not arranged such that they would simultaneously land within the relevant receptive fields.

Participants performed a change detection task with colored squares presented either across or within hemifields. In addition, items were presented either closer together or farther apart. Multiple low-level receptive field size estimates were used to constrain where the items were placed. Here, we used V4 as a proxy for the visual area most likely to be directly involved in representing the color of the squares (Wandell et al., 2007). All items were placed at an eccentricity of 11.4° . At this distance, V4 receptive fields have been estimated to be $\sim 7^\circ$ to $\sim 8.5^\circ$ (Gattass et al., 1988; Smith et al., 2001; Winawer et al., 2010). In the *far condition*, items were placed far enough apart (9.8°) so there would be few cases of multiple items landing within a V4 receptive field. In the *close condition*, items were placed close enough (6.3°) so there would likely be many instances of multiple items landing within the same V4 receptive fields. Finally, we assume that these smaller receptive fields are also lateralized, in which case V4 receptive fields do not span between the left/right hemifield (Gattass et al., 1988) (Figure 4a, note the absence of receptive fields along the vertical midline), but do span the upper/lower hemifields (Figure 4a, note the presence of receptive fields along the horizontal midline). Thus, if receptive field interference caused the pattern of results observed in Experiments 1a and 1b, there should be a hemifield effect with colors in the close condition, but not the far condition.

Method

Participants. Twenty participants were recruited from the Harvard University community with reported normal or corrected-to-normal Vision.

Stimuli. The stimuli were $12 \times 2.95^\circ \times 2.95^\circ$ color squares evenly distributed around 40% of the circle in $CIE L^*a^*b$ color space centered at $L = 54$, $a = 18$, $b = -8$, with a radius of 59. Only a portion of the color space was used to increase the difficulty and avoid ceiling performance. All items were on a gray background, which had a luminance of 65.1 cd/m^2 .

Procedure. As was the case with Experiment 1, participants performed a change detection task in which they had to detect

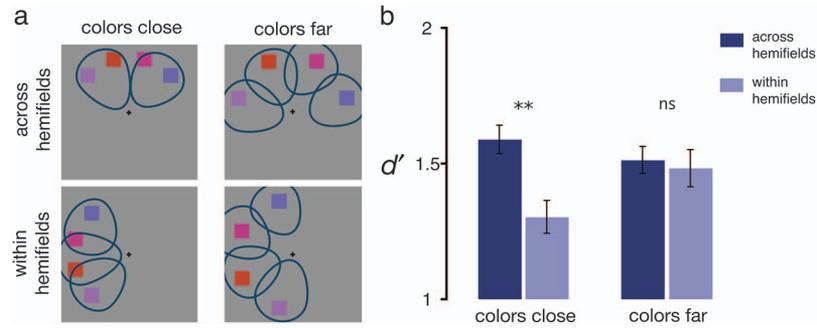


Figure 4. (a) Idealized schematic of receptive field size in relation to the four conditions of Experiment 2. (b) Results from Experiment 2. Twenty participants performed the experiment. Asterisks indicate significant differences ($p < .05$), within-subjects two-tailed t tests. Error bars reflect within-subject SEM (Cousineau, 2005).

changes between two successive displays. The initial display appeared for 800 ms, followed by a 500 ms blank fixation screen, and then the second, probe display, which remained until participants responded using the keyboard. Visual feedback was given immediately. Changes occurred on half the trials with changes occurring at each location of each configuration equally often. On change trials, one color was replaced by another color that was 52° away on the color wheel. There were 10 practice trials and 320 experimental trials that were divided into two blocks: the close block and the far block. Eye-tracking calibration sequences occurred before the practice trials, the experimental block, and halfway through the experimental block.

Results and Discussion

As predicted, a hemifield effect was found when the colors were closer together (across $d' = 1.59$, $SEM = 0.05$; within $d' = 1.30$, $SEM = 0.06$: $t(19) = 3.98$, $p < .001$), but not when they were farther apart (across $d' = 1.51$, $SEM = 0.05$; within $d' = 1.48$, $SEM = 0.06$: $t(19) = 0.39$, $p = .70$; Figure 4). To determine if the size of the hemifield effect was larger when the colors were closer than when they were farther apart, we performed a within subjects t test on the differences between across and within hemifield presentation for the two conditions. This revealed a significant difference in the size of the hemifield effect as a function of item spacing, $t(19) = 2.16$, $p < .05$.

Although we only see a hemifield effect for colors when items are presented close together, the perceptual interference measured in this experiment should be distinguished from visual crowding. Crowding is the impaired ability to identify a stimulus in the periphery when it is flanked by nearby stimuli, and typically occurs when the flanking objects are within 0.5 x the target eccentricity (e.g., if the target is 10° away, crowding occurs from flankers within 5° ; Pelli, 2008; Whitney & Levi, 2011).

To verify that items in the close condition were not crowded, we replicated the experiment with one change: on every trial, one item was cued and participants only attended to and remembered that item.¹ Only the cued item could ever change. In this experiment, performance was near ceiling in each of the four conditions ($d' > 3.20$; accuracy $> 97\%$), suggesting that the items were not crowded in the traditional sense. Furthermore, there was no hemi-

field effect in either the far (across $d' = 3.24$, $SEM = 0.05$; within $d' = 3.29$, $SEM = 0.04$: $t(7) = 0.62$, $p = 0.56$) or close (across $d' = 3.29$, $SEM = 0.05$; within $d' = 3.37$, $SEM = 0.06$: $t(7) = 0.80$, $p = 0.45$) condition, with no interaction between the conditions, $t(7) = 0.99$, $p = 0.35$. To be sure these null findings were not simply because of differences in the number of participants between the main experiment and the cue condition, we calculated the Bayes factor for the hemifield effect in the close condition and the interaction between the hemifield effects in the close and far conditions, both of which were significant in the main experiment. In both cases, the Bayes factor was below .33 (close condition = 0.15; interaction = 0.31), which supports our acceptance of the null hypothesis (Dienes, 2011). This suggests that the hemifield effect observed in the close condition was not caused by visual crowding because participants' performance was at ceiling in each of the different conditions. Instead, this suggests that the hemifield effect was the result of competition between items when attention was divided across those items. This is consistent with several physiological results showing that multiple items within the same receptive field interfere with one another if focal attention is not deployed (Beck & Kastner, 2005; Desimone & Duncan, 1995; Scaf & Beck, 2010; Zoccolan et al., 2005). Similar to those findings, we find that attending to the cued items diminishes the competition between items and in this case, eliminates the hemifield effect.

Experiment 3

In the previous experiments, we examined how receptive fields affect perception by measuring the behavioral hemifield effect.

¹ *Color cue condition*: Eight new participants performed 192 trials in the cue condition in which a single item was cued on each trial. In this case, a white, 8.7° long line was drawn from the fixation point toward an item. The line appeared 500 ms before the first display and remained the screen throughout the remainder of the trial. The only item that could change was the cued item and the cued item changed on half the trials. One of the two middle locations, where crowding would be maximal, was always cued, while the two outer locations were never cued. Each of the two middle items were cued equally often in each of the four configurations. Besides the presence of the cue and the fact that only the cued item could change, every aspect of the cue condition was the same as noncued condition (i.e., number of trials, presentation duration, etc.).

However, if receptive field competition caused these spatial interference patterns, we should also find spacing effects within a hemifield, particularly for low-level stimuli. Because low-level receptive fields are relatively small, it is possible to move multiple low-level items in and out of the same receptive fields by moving them closer together or farther apart (Figure 5a). Thus, behavioral performance should be worse when low-level items are closer together. Because IT receptive fields are relatively large (greater than $\sim 20^\circ$ in the periphery; Op de Beeck & Vogels, 2000; Distler et al., 1993), moving high-level items closer or farther apart within a hemifield should have little effect on competition between items since the items will always remain within the receptive field regardless of spacing. Thus, behavioral performance for high-level items should be less sensitive to changes in spacing within a hemifield.

We used estimates of receptive fields in the literature to constrain where the items were placed. In the close condition, items were close enough (5.5°) that multiple items could fall within a typical V4 receptive field ($\sim 6^\circ$ to $\sim 7.5^\circ$, Gattass et al., 1988; Smith et al., 2001; Winawer et al., 2010). In the far condition, items were placed far enough apart (12.4°) that a typical V4 receptive field would encompass only one item, but close enough that a large IT receptive field, would encompass multiple objects. Thus, changing the spacing of the items should affect the input to low-level, but not high-level, receptive fields.

Method

Participants. Twenty participants were recruited from the Harvard University community with reported normal or corrected-to-normal vision.

Stimuli. Stimuli were real world objects and colored squares. The real world objects were grayscale and matched within category for luminance and contrast. The colored squares were the same ones used in Experiment 2. All stimuli were placed within a $2.95^\circ \times 2.95^\circ$ square. In all conditions, three items were presented 10° away from the fixation point. The center-to-center distance between adjacent stimuli was 5.5° in the close condition and 12.4° in the far condition. All items were always presented within a single hemifield on a given trial with items being presented in the left and right hemifields equally often in all four conditions. In the

close condition, all items were drawn in the top visual quadrants, with the middle item in the center of the quadrant (equal distance from the vertical and horizontal meridians). In the far condition, the center item was placed on the horizontal meridian with one other item placed in the top and bottom visual quadrants. All items were on a gray background, which had a luminance of 65.1 cd/m^2 .

Procedure. Participants once again performed a change detection task with three items that always appeared within a visual hemifield. The initial display appeared for 800 ms, followed by a 500 ms blank fixation screen, and then the second, probe display, which remained until participants responded using the keyboard. Visual feedback was given immediately. There were four conditions: colors far, colors close, objects far, and objects close. Each of these conditions was performed in a separate block. Within each block there were 15 practice and 80 experimental trials. The order in which these blocks occurred was counterbalanced using a Latin square design. Changes occurred on half of the trials in each block. The location of the change item was randomly selected on every trial from the three possible locations. On color change trials, one color was replaced by another color that was 39° away on the color wheel. We limited the number of degrees a color could change on the color wheel simply to avoid ceiling performance. On object change trials, one object was replaced by a new, randomly selected object. Eye-tracking calibration sequences occurred before the practice and experimental trials of the first block, as well as before the experimental trials of the third block.

Results and Discussion

As predicted, performance with the colored squares was affected by moving the items closer or farther apart (far $d' = 1.32$, $SEM = 0.05$; close $d' = 1.00$, $SEM = 0.06$; $t(19) = 4.26$, $p < .001$; Figure 5b). In contrast, moving the objects in the same manner did not affect performance (far $d' = 1.38$, $SEM = 0.07$; close $d' = 1.42$, $SEM = 0.07$; $t(19) = 0.34$, $p = 0.73$). In addition, there was a significant interaction between object type and the size of the spacing effect, $t(19) = 2.67$, $p < .05$.

While these results are consistent with differences in receptive field sizes, it is again important to ask if these effects can be explained by crowding. We addressed this by repeating the experiment with a cue that informed participants of where the change

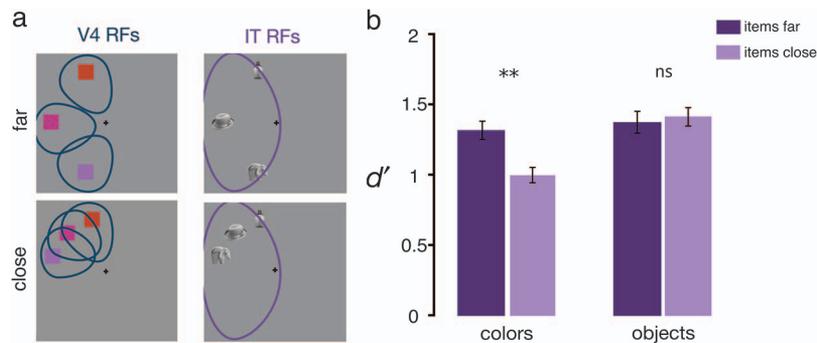


Figure 5. (a) Idealized schematic of receptive field sizes in low-level and high-level regions in relation to the four conditions of Experiment 3. (b) Results from Experiment 3. Asterisks indicate significant differences ($p < .05$), within-subjects two-tailed t tests. Error bars reflect within-subject SEM (Cousineau, 2005). RFs = receptive fields; IT = inferotemporal.

might occur.² Once again, we found that performance in all conditions was near ceiling ($d' > 3.10$; accuracy $> 94\%$), when a single item was cued. Furthermore, the presence of the cue eliminated all differences between the close and far conditions (colors far $d' = 3.14$, $SEM = 0.07$; colors close $d' = 3.17$, $SEM = 0.13$: $t(7) = 0.30$, $p = 0.78$; objects far $d' = 3.11$, $SEM = 0.08$; objects close $d' = 3.12$, $SEM = 0.10$: $t(7) = 0.12$, $p = .91$; interaction: $t(7) = 0.17$, $p = .87$). Furthermore, we again calculated the Bayes factor for the spacing effect in the color condition and the interaction between the spacing effect in the color and object conditions, both of which were significant in the main experiment. The Bayes factor was 0.36 for the color condition and 0.32 for the interaction, which lends support for our acceptance of the null hypothesis. These results suggest that these spatial interference effects occur when participants attempt to simultaneously encode multiple items.

Experiment 4

Finally, we asked if this spatial interference occurs during sensory encoding or memory maintenance. To answer this, we asked if a hemifield effect is observed with high-level stimuli when items are presented either simultaneously or sequentially. When items are presented simultaneously, they land within the same receptive fields and compete with one another. When items are presented sequentially, they never land within the same receptive fields simultaneously because only one item is on the display at any given time. Thus, if the hemifield effect is caused by interference within receptive fields only during encoding, there should only be an effect during the simultaneous condition. However, because the storage demands are comparable for both conditions, if this competition arises in part because of storage limitations, spatial interference should be observed for both simultaneous and sequential displays. Items were initially presented in both conditions for either 200 or 800 ms. Varying the encoding time enables us to compare the simultaneous and sequential conditions, matched either in terms of time per item or total presentation time.

Method

Participants. Two groups of 18 participants performed the experiment at the two different encoding intervals: 200 ms/800 ms. Both groups were comprised of participants that were recruited from the Harvard University community and reported normal or corrected-to-normal vision.

Stimuli. The object stimuli from Experiment 1 were used for this experiment; faces and scenes were not used. All items were on a gray background, which had a luminance of 68.5 cd/m².

Procedure. In both the simultaneous and sequential conditions, stimuli were presented either across or within the visual hemifields in the same locations as those in Experiment 1. The simultaneous condition was similar to Experiment 1 with following exceptions: (a) there were four different retention intervals to match the sequential condition (see below) and (b) only one item was presented during the second display (i.e., probe display) rather than four items. Participants indicated whether that particular item changed using the keyboard. Changes always involved one object being replaced by a new object that was not present among the four target items.

In the sequential condition, 500 ms would always pass between the presentation of the last encoding item and the presentation of the probe display, which, like the simultaneous condition, only contained one item. However, the amount of time that would pass between the initial presentation of an item and its subsequent probe would vary. Consider the 800 ms condition. In the sequential block, each individual item was shown one at a time for 800 ms. If the first item presented was probed, 2,900 ms would pass between the offset of that item and the probe display. If the last item was probed, 500 ms would pass. Each temporal position of the initial presentation (first, second, third, or fourth) was probed an equal number of times. Thus, in the 800 ms/item condition, the time between presentation and probe were 2,900, 2,100, 1,300, and 500 ms. In the 200 ms/item condition, the retention intervals were 1,100, 900, 700, and 500 ms. In addition, within each of the four different spatial configurations of items (i.e., across-above, across-below, within-left, or within-right), each of the four possible locations was probed equally often. Thus, subjects could not use either temporal or spatial information to predict which item would be probed.

In the simultaneous condition, four items were presented for 200 ms or 800 ms, followed by a variable blank delay that matched the retention intervals for the sequential displays (200 ms condition: 500 ms, 700 ms, 900 ms, or 1,100 ms; 800 ms condition: 500 ms, 1,300 ms, 1,900 ms, or 2,700 ms; Figure 6a). Varying the retention interval for the simultaneous condition enabled us to ensure that the average retention duration for the probed item was matched between the conditions.

Results and Discussion

For both retention intervals, there was a hemifield effect in the simultaneous condition (200 ms across $d' = 1.76$, $SEM = 0.07$; 200 ms within $d' = 1.34$, $SEM = 0.09$: $t(17) = 4.11$, $p < 0.001$; 800 ms across $d' = 2.05$, $SEM = 0.09$; 800 ms within $d' = 1.68$, $SEM = 0.09$: $t(17) = 2.70$, $p < 0.05$), but not in the sequential condition (200 ms across $d' = 1.95$, $SEM = 0.06$; 200 ms within $d' = 1.85$, $SEM = 0.09$: $t(7) = 1.17$, $p = 0.26$; 800 ms across $d' = 2.34$, $SEM = 0.08$; 800 ms within $d' = 2.28$, $SEM = 0.10$: $t(17) = 0.52$, $p = 0.61$; Figure 6b). We then performed a 2×2 ANOVA on presentation (simultaneous vs. sequential) and stimulus location (across vs. within hemifield) with the presentation durations (200 ms vs. 800 ms) entered as random factors because they were carried out with two sets of participants. In this case, the interaction between stimulus presentation and display configuration was significant ($F(1, 35) = 6.68$, $p < 0.05$), while the interaction between stimulus presentation, display configuration, and display duration was not, $F(1, 35) = 0.01$, $p > .99$. These results suggest that stimuli appear to compete with one another primarily during encoding, and not during memory maintenance. It is also worth

² *Object cue condition:* Eight new participants performed 192 trials in the cue condition in which a single item was cued on each trial. One item was cued by a white, 8.7° long line that was drawn from the fixation point toward an item. The line appeared 500 ms before the first display and remained the screen throughout the remainder of the trial. Besides the presence of the cue and the fact that only the cued item could change, every aspect of the cue condition was the same as noncued condition (i.e., number of trials, presentation duration, etc.). Only the center item was ever cued and thus only the center item could ever change.

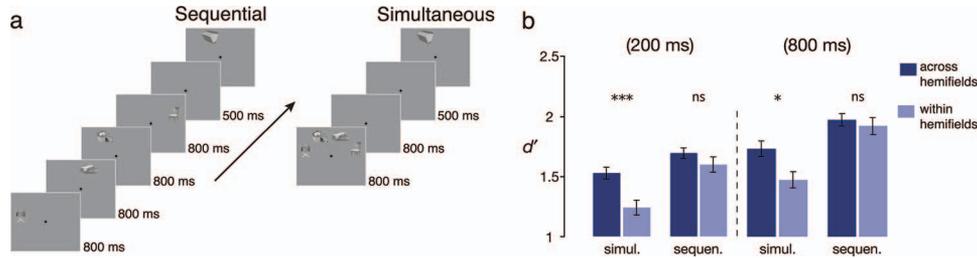


Figure 6. (a) Example of the sequential condition (left panel) and the simultaneous condition (right panel) when presented for 800 ms/item, and matched in terms of the storage duration of the probed item. In this example, the third item in the sequential trial is subsequently tested. Thus, to match the amount of time between the initial presentation of this item and the test display, the retention interval in the simultaneous trial is adjusted. This was done across all possible retention intervals in both the 200 ms/item and 800 ms/item conditions. (b) Results from Experiment 4 for the two presentations durations (200 ms/item and 800 ms/item) for both simultaneous and sequential presentations. Eighteen participants performed the experiment at the 200 ms duration and another 18 participants performed the experiment at the 800 ms duration. Asterisks indicate significant differences ($p < .05$), within-subjects two-tailed t tests. Error bars reflect within-subject *SEM* (Cousineau, 2005).

noting that there was a main effect of presentation format with performance on the sequential conditions being greater than performance on the simultaneous conditions, $F(1, 35) = 20.44$, $p < .001$. This is consistent with previous results showing that presenting items sequentially improves visual short-term memory (STM) performance (Ihssen et al., 2010; Shapiro & Miller, 2011). This further supports the notion that interference between items occurs during encoding and not maintenance.

Of interest to the authors, this finding differs from a previous result, which found a hemifield effect with both simultaneous and sequential presentations (Umemoto et al., 2010). One possible reason for this discrepancy may stem from the fact that this previous finding required participants to remember the orientation of target items, while we had participants encode object identity. It is possible that these tasks rely on different neural networks and there are differences in the hemifield effect (i.e., competition during encoding or maintenance) depending on the relevant cortical region. The task used by Umemoto and colleagues potentially relied more heavily on the dorsal visual pathway, where representations of location and spatial information are more precisely encoded (Ungerleider & Mishkin, 1982). The task reported here, meanwhile, likely relied on the ventral pathway, because object identities had to be processed, and spatial information was not a critical factor. Thus, it is possible that hemifield effects stemming from the ventral pathway occur primarily during encoding because items can only compete with one another in receptive fields during the initial encoding process. Meanwhile, hemifield effects stemming from the dorsal pathway likely arise for different reasons (i.e., asymmetries in spatial suppressive effects during attentional selection, see below for extended discussion) and can constrain visual cognition during both encoding and maintenance. Of course, this explanation is merely speculation and future research will be needed to examine this issue.

Discussion

We asked if receptive field competition limits the ability to encode multiple items simultaneously. To test this, we made

stimulus-specific predictions for the degree of interference that should be observed between items at different locations (e.g., different predictions for faces vs. color patches), based on the known size and spatial distribution of receptive fields. We found evidence consistent with the idea that receptive fields in the ventral visual stream limit the ability to perceive multiple items simultaneously. Specifically, both high-level and low-level items show greater interference within a hemifield than across hemifields (Experiments 1 and 2), with interference operating over a long-range for high-level items, and short range for low-level items (Experiment 1–3). These interference effects disappear when attention is directed to one particular item (Experiments 2 and 3) and appear to reflect limits on the ability to simultaneously encode multiple items (Experiment 4).

It should be stressed that the different spatial interference effects we observe are not necessarily inherent to certain stimulus categories (e.g., faces). Instead, it is likely that both the stimuli and the task being performed will determine the interference between items. Consider a hypothetical change detection experiment in which every face was a different color, such that whenever a face changes, the color changes. In this case, because participants could perform the task by noticing changing colors, the critical receptive fields may be in a region like V4, and so we might only observe short-range interference. Meanwhile, if the task relied on processing facial identity, IT would become critical and we would then observe long-range interference. Thus, isolating higher levels of representation requires constructing stimuli in which low-level feature differences are minimized.

Anatomical Constraints of Both Visual Streams

Considerable research has investigated the benefits of presenting information across the two visual hemifields. Typically, visual hemifield effects are attributed to separate pools of attention in each cerebral hemisphere (Luck et al., 1989). Many tasks demonstrating attentional hemifield effects have significant spatial demands (Alvarez & Cavanagh, 2005; Buschman et al., 2011; Delvenne, 2005; Umemoto et al., 2010). Oftentimes, if those spatial

demands are removed, the hemifield effects disappear (Alvarez et al., 2012; Awh & Pashler, 2000; Delvenne & Holt, 2012). The need for spatial demands to elicit hemifield effects is potentially because of the surround suppression mechanisms associated with attention. When an object is attended to, the regions surrounding that object are actively suppressed and inhibited (Hopf et al., 2006; Störmer & Alvarez, 2014). It may be the case that attending to an object in one hemifield suppresses other information within that hemifield, but does not suppress information in the other hemifield; thus, a benefit for presenting stimuli across hemifields but not within hemifields. The reason there is this strict hemifield divide is that a variety of attentional networks are located in the parietal lobe, which has been shown to have weaker communication between the two cerebral hemispheres than the ventral stream (Ungerleider & Mishkin, 1982). With this fact in mind, it is natural to think that hemifield effects will not be found with tasks that rely on the ventral pathway since (a) the tasks we used have no spatial demands and (b) the two cerebral hemispheres are relatively well-connected compared with the dorsal stream. Counter to this prediction, however, we still found consistent hemifield effects with these identity-based tasks that are likely supported by the ventral pathway. We suggest that these particular effects are because of competition within ventral stream receptive fields (Clevenger & Beck, 2014; Torralbo & Beck, 2008). Thus, we propose that both visual pathways impose anatomical constraints on visual cognition in different ways. Under this view, the dorsal stream constrains cognition because attentional suppressive mechanisms that are implemented in the two, separate hemispheres of the parietal lobe. The ventral stream, meanwhile, constrains cognition from competition within receptive fields.

The Nature of Receptive Field Interference

When multiple items land within a receptive field, the neuronal output will be the average of the neuron's response to the constituent items when those items are presented in isolation (Zoccolan et al., 2005). This averaging process will result in representations of those particular items that are less precise. What exactly happens to the representations of the individual items when they land within a single receptive field? One possibility is that the features of the different items will be pooled together, resulting in a representation based largely on the items' summary statistics (Alvarez, 2011). Another possibility is that this averaging process will lead to an increase in binding errors. In support of this, it was recently found that presenting multiple colors close to one another in a working memory task increased the number of binding errors between a stimulus and its location (Emrich & Ferber, 2012). Isolating the exact nature of the interference between items will require more precise measures of the content of representations that are formed when multiple items are encoded simultaneously.

Relation to Crowding

The spatial interference patterns reported here resemble those seen in visual crowding. Crowding is an impaired ability to recognize peripheral items among clutter and is thought to occur because of feature integration over an inappropriately large area of peripheral space. Traditionally, this overintegration is thought to occur if objects are presented too close to one another in early

visual cortex (Pelli, 2008; Whitney & Levi, 2011). When multiple items land within early visual integration/receptive fields, the integration of features prevents the identification of target items among distractors.

While it is unclear if the results reported here are because of feature integration or an increase in binding errors, we suggest that these effects represent a type of "later-stage crowding." When attention is divided across multiple items, those items "crowd" one another and decrease performance on a perceptual task. On this account, crowding appears to be a ubiquitous process that affects visual processing at multiple levels of the visual hierarchy. However, the critical difference between classical crowding and the later-stage crowding effects observed here is the role of attention. In classical crowding, while precuing attention to the target may slightly reduce the critical distance between the target and distractors (Yeshurun & Rashal, 2010; but see Scolari et al., 2007), it does not completely eliminate crowding (Whitney & Levi, 2011). However, in Experiments 2 and 3 in this study, we found that participants had no problem identifying the target regardless of what kind of item it was (i.e., color patch or object) when it was precued, and performance was at ceiling in all precued conditions. These results suggest that these items are not crowded in the traditional sense and also show that competition between items that land within large receptive fields (relative to, say, V1) is resolved by attention. This is likely caused by attention eliminating unwanted, unattended inputs to high-level receptive fields (Desimone & Duncan, 1995), presumably by modulating input from or activity in lower-level regions.

Relation to Models of Attention and Working Memory

Focusing on the underlying neural architecture of the visual system provides critical constraints on various models of cognition. For example, prominent models of visual attention and working memory propose that attention and memory are limited by either a fixed number of discrete representations (Buschman et al., 2011; Cowan, 2001; Zhang & Luck, 2008) or a limited amount of cognitive resources (Alvarez & Franconeri, 2007; Bays & Husain, 2008; Holcombe & Chen, 2012). However, in their current form, neither model would explicitly predict the pattern of spatial interference effects observed here. These results do fit well with models in which perceptual limitations are based on competition within feature maps (Franconeri et al., 2013; Reddy & VanRullen, 2007; Scaf & Beck, 2010; Scaf, Torralbo, Tapia, & Beck, 2013; VanRullen et al., 2005). However, because we found interference effects during simultaneous presentation, but not sequential presentation, it appears that map-interference plays a role in perceptual encoding more so than for memory maintenance.

Conclusion

Based on the current results, we propose that cognitive models of attention and working memory need to account for stimulus-specific differences in perceptual encoding. Otherwise, such models risk misattributing differences in performance across stimuli to higher-level cognitive processes. Overall, we suggest that the present results demonstrate how psychologists can benefit from considering findings from neuroscience. Abstract models of cognition can be clarified and refined by focusing on neural imple-

mentations (Alvarez et al., 2012; Torralbo & Beck, 2008). In addition, findings from neuroscience can also lead to the formation of novel psychological hypotheses that would not otherwise be developed and tested.

References

- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8, 457–464. <http://dx.doi.org/10.1016/j.tics.2004.08.011>
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15, 122–131. <http://dx.doi.org/10.1016/j.tics.2011.01.003>
- Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science*, 16, 637–643. <http://dx.doi.org/10.1111/j.1467-9280.2005.01587.x>
- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, 7, 1–10. <http://dx.doi.org/10.1167/7.13.14>
- Alvarez, G. A., Gill, J., & Cavanagh, P. (2012). Anatomical constraints on attention: Hemifield independence is a signature of multifocal spatial selection. *Journal of Vision*, 12, 9. <http://dx.doi.org/10.1167/12.5.9>
- Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 834–846. <http://dx.doi.org/10.1037/0096-1523.26.2.834>
- Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology*, 8, 234–238. [http://dx.doi.org/10.1016/S0959-4388\(98\)80145-1](http://dx.doi.org/10.1016/S0959-4388(98)80145-1)
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321, 851–854. <http://dx.doi.org/10.1126/science.1158023>
- Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience*, 8, 1110–1116. <http://dx.doi.org/10.1038/nm1501>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436. <http://dx.doi.org/10.1163/156856897X00357>
- Broadbent, D. (1958). *Perception and communication*. London: Pergamon Press. <http://dx.doi.org/10.1037/10037-000>
- Buschman, T. J., Siegel, M., Roy, J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 11252–11255. <http://dx.doi.org/10.1073/pnas.1104666108>
- Clevenger, J., & Beck, D. M. (2014). Refining the resource model: Cortical competition could explain hemifield independence. *Visual Cognition*, 22, 1022–1026. <http://dx.doi.org/10.1080/13506285.2014.960668>
- Cohen, M. A., Konkle, T., Rhee, J. Y., Nakayama, K., & Alvarez, G. A. (2014). Processing multiple visual objects is limited by overlap in neural channels. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 8955–8960. <http://dx.doi.org/10.1073/pnas.1317860111>
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1, 42–45.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–114. <http://dx.doi.org/10.1017/S0140525X01003922>
- Delvenne, J.-F. (2005). The capacity of visual short-term memory within and between hemifields. *Cognition*, 96, B79–B88. <http://dx.doi.org/10.1016/j.cognition.2004.12.007>
- Delvenne, J.-F., & Holt, J. L. (2012). Splitting attention across the two visual fields in visual short-term memory. *Cognition*, 122, 258–263. <http://dx.doi.org/10.1016/j.cognition.2011.10.015>
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222. <http://dx.doi.org/10.1146/annurev.ne.18.030195.001205>
- Deutsch, J. A., & Deutsch, D. (1963). Some theoretical considerations. *Psychological Review*, 70, 80–90. <http://dx.doi.org/10.1037/h0039515>
- Dienes, Z. (2011). Versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6, 274–290. <http://dx.doi.org/10.1177/1745691611406920>
- Distler, C., Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1993). Cortical connections of inferior temporal area TEO in macaque monkeys. *The Journal of Comparative Neurology*, 334, 125–150. <http://dx.doi.org/10.1002/cne.903340111>
- Emrich, S. M., & Ferber, S. (2012). Competition increases binding errors in visual working memory. *Journal of Vision*, 12, 1–16. <http://dx.doi.org/10.1167/12.4.12>
- Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences*, 17, 134–141. <http://dx.doi.org/10.1016/j.tics.2013.01.010>
- Gallace, A., Tan, H. Z., & Spence, C. (2006). The failure to detect tactile change: A tactile analogue of visual change blindness. *Psychonomic Bulletin & Review*, 13, 300–303. <http://dx.doi.org/10.3758/BF03193847>
- Gattass, R., Sousa, A. P., & Gross, C. G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *The Journal of Neuroscience*, 8, 1831–1845.
- Golomb, J. D., & Kanwisher, N. (2012). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex*, 22, 2794–2810. <http://dx.doi.org/10.1093/cercor/bhr357>
- Graziano, M. S., & Gross, C. G. (1993). A bimodal map of space: Somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Experimental Brain Research*, 97, 96–109. <http://dx.doi.org/10.1007/BF00228820>
- Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, 36, 791–804. [http://dx.doi.org/10.1016/S0896-6273\(02\)01091-7](http://dx.doi.org/10.1016/S0896-6273(02)01091-7)
- Holcombe, A. O., & Chen, W. Y. (2012). Exhausting attentional tracking resources with a single fast-moving object. *Cognition*, 123, 218–228. <http://dx.doi.org/10.1016/j.cognition.2011.10.003>
- Hopf, J.-M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H.-J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 1053–1058. <http://dx.doi.org/10.1073/pnas.0507746103>
- Ihssen, N., Linden, D. E. J., & Shapiro, K. L. (2010). Improving visual short-term memory by sequencing the stimulus array. *Psychonomic Bulletin & Review*, 17, 680–686. <http://dx.doi.org/10.3758/PBR.17.5.680>
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Knudsen, E. I. (1982). Auditory and visual maps of space in the optic tectum of the owl. *The Journal of Neuroscience*, 2, 1177–1194.
- Kravitz, D. J., Vinson, L. D., & Baker, C. I. (2008). How position dependent is visual object recognition? *Trends in Cognitive Sciences*, 12, 114–122. <http://dx.doi.org/10.1016/j.tics.2007.12.006>
- Luck, S. J., Hillyard, S. A., Mangun, G. R., & Gazzaniga, M. S. (1989). Independent hemispheric attentional systems mediate visual search in split-brain patients. *Nature*, 342, 543–545. <http://dx.doi.org/10.1038/342543a0>
- Mance, I., Becker, M. W., & Liu, T. (2012). Parallel consolidation of simple features into visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 429–438. <http://dx.doi.org/10.1037/a0023925>
- Motter, B. C. (2009). Central V4 receptive fields are scaled by the V1 cortical magnification and correspond to a constant-sized sampling of

- the V1 surface. *The Journal of Neuroscience*, 29, 5749–5757. <http://dx.doi.org/10.1523/JNEUROSCI.4496-08.2009>
- Op De Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *The Journal of Comparative Neurology*, 426, 505–518. [http://dx.doi.org/10.1002/1096-9861\(20001030\)426:4<505::AID-CNE1>3.0.CO;2-M](http://dx.doi.org/10.1002/1096-9861(20001030)426:4<505::AID-CNE1>3.0.CO;2-M)
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. <http://dx.doi.org/10.1163/156856897X00366>
- Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology*, 18, 445–451. <http://dx.doi.org/10.1016/j.conb.2008.09.008>
- Reddy, L., & VanRullen, R. (2007). Spacing affects some but not all visual searches: Implications for theories of attention and crowding. *Journal of Vision*, 7, 1–17. <http://dx.doi.org/10.1167/7.2.3>
- Scalf, P. E., & Beck, D. M. (2010). Competition in visual cortex impedes attention to multiple items. *The Journal of Neuroscience*, 30, 161–169. <http://dx.doi.org/10.1523/JNEUROSCI.4207-09.2010>
- Scalf, P. E., Torralbo, A., Tapia, E., & Beck, D. M. (2013). Competition explains limited attention and perceptual resources: Implications for perceptual load and dilution theories. *Frontiers in Psychology*, 4, 243. <http://dx.doi.org/10.3389/fpsyg.2013.00243>
- Scolari, M., Kohnen, A., Barton, B., & Awh, E. (2007). Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays? *Journal of Vision*, 7, 1–23. <http://dx.doi.org/10.1167/7.2.7>
- Shapiro, K. L., & Miller, C. E. (2011). The role of biased competition in visual short-term memory. *Neuropsychologia*, 49, 1506–1517. <http://dx.doi.org/10.1016/j.neuropsychologia.2011.02.017>
- Smith, A. T., Singh, K. D., Williams, A. L., & Greenlee, M. W. (2001). Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cerebral Cortex*, 11, 1182–1190. <http://dx.doi.org/10.1093/cercor/11.12.1182>
- Störmer, V. S., & Alvarez, G. A. (2014). Feature-based attention elicits surround suppression in feature space. *Current Biology*, 24, 1985–1988. <http://dx.doi.org/10.1016/j.cub.2014.07.030>
- Torralbo, A., & Beck, D. M. (2008). Perceptual-load-induced selection as a result of local competitive interactions in visual cortex. *Psychological Science*, 19, 1045–1050. <http://dx.doi.org/10.1111/j.1467-9280.2008.02197.x>
- Umemoto, A., Drew, T., Ester, E. F., & Awh, E. (2010). A bilateral advantage for storage in visual working memory. *Cognition*, 117, 69–79. <http://dx.doi.org/10.1016/j.cognition.2010.07.001>
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549–586). Cambridge, MA: MIT Press.
- VanRullen, R., Reddy, L., & Fei-Fei, L. (2005). Binding is a local problem for natural objects and scenes. *Vision Research*, 45, 3133–3144. <http://dx.doi.org/10.1016/j.visres.2005.05.012>
- Vitevitch, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 333–342.
- Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56, 366–383. <http://dx.doi.org/10.1016/j.neuron.2007.10.012>
- Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, 15, 160–168. <http://dx.doi.org/10.1016/j.tics.2011.02.005>
- Wilson, D. A. (2001). Receptive fields in the rat piriform cortex. *Chemical Senses*, 26, 577–584. <http://dx.doi.org/10.1093/chemse/26.5.577>
- Winawer, J., Horiguchi, H., Sayres, R. A., Amano, K., & Wandell, B. A. (2010). Mapping hV4 and ventral occipital cortex: The venous eclipse. *Journal of Vision*, 10, 1–22. <http://dx.doi.org/10.1167/10.5.1>
- Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision*, 10, 16. <http://dx.doi.org/10.1167/10.10.16>
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233–235. <http://dx.doi.org/10.1038/nature06860>
- Zoccolan, D., Cox, D. D., & DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *The Journal of Neuroscience*, 25, 8150–8164. <http://dx.doi.org/10.1523/JNEUROSCI.2058-05.2005>

Received February 12, 2015

Revision received June 5, 2015

Accepted June 8, 2015 ■