The role of spatial attention in crowding and feature binding

Bahiyya Kewan-Khalayly

Department of Special Education, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel ✉

Amit Yashar

Department of Special Education, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel 🏠✉

Crowding refers to the failure to identify a peripheral object due to nearby objects (flankers). A hallmark of crowding is inner–outer asymmetry; that is, the outer flanker (more peripheral) produces stronger interference than the inner one. Here, by manipulating attention, we tested the predictions of two competing accounts: the attentional account, which predicts a positive attentional effect on the inner–outer asymmetry (i.e., attention to the outer flanker will increase asymmetry) and the receptive field size account, which predicts a negative attentional effect. In Experiment 1, observers estimated a Gabor target orientation. A peripheral pre-cue drew attention to one of three locations: target, inner flanker, or outer flanker. Probabilistic mixture modeling demonstrated asymmetry by showing that observers often misreported the outer-flanker orientation as the target. Interestingly, the outer cue led to a higher misreport rate of the outer flanker, and the inner cue led to a lower misreport rate of the outer flanker. Experiment 2 tested the effect of crowding and attention on incoherent object reports (i.e., binding errors, reporting the tilt of one presented item with the color of another item). In each trial, observers estimated both the tilt and color of the target. Attention merely increased coherent target reports, but not coherent flanker reports. The results suggest that the locus of spatial attention plays an essential role in crowding, as well as inner–outer asymmetry, and demonstrate that crowding and feature binding are closely related. However, our findings are inconsistent with the view that covert attention automatically binds features together.

Introduction

In vision, the spacing of objects can fundamentally limit object recognition. Objects that are too close together can become indistinguishable (cluttered), a phenomenon known as "crowding" (Pelli, 2008; Whitney & Levi, 2011). Crowding hinders the identification of various basic stimuli, such as letters (Bouma, 1970) and faces (Strasburger, Rentschler, & Jüttner, 2011), and impairs essential perceptual tasks such as reading (Whitney & Levi, 2011) and face recognition (Strasburger et al., 2011). Crowding plays a critical role in deficits such as macular degeneration (Wallace, Chung, & Tjan, 2017), amblyopia (Song, Levi, & Pelli, 2014), and dyslexia (Gori & Facoetti, 2015). Recently, it has been shown that crowding errors directly reflect binding errors (i.e., reporting a feature of one item as belonging to another item) (Yashar, Wu, Chen, & Carrasco, 2019). Thus, investigating crowding and the means to reduce its disruptive effect has important implications for understanding object recognition and has the potential for clinical contribution (Levi, 2008).

The critical spacing of crowding—that is, the minimal space between the target and the flankers that permits performance similar to when no flankers are presented—scales with target eccentricity (Bouma, 1970; Pelli, Palomares, & Majaj, 2004). Although spatial attention (Yeshurun & Rashal, 2010) and training (perceptual learning) (Chung, 2007; Hussain, Webb, Astle, & McGraw, 2012; Yashar, Chen, & Carrasco, 2015; Zhu, Fan, & Fang, 2016) can reduce the critical spacing, their effect on the interference is limited, and in a typical crowded display the critical spacing is often 0.3 to 0.5 of eccentricity (Levi, 2008; Pelli et al., 2004).

A hallmark of crowding is inner–outer asymmetry; that is, in a radial display, an outer flanker (more peripheral) produces stronger interference than an inner one (closer to the fovea) (Banks, Bachrach, & Larson, 1977; Bouma, 1970; Chaney, Fischer, & Whitney, 2014; Dayan & Solomon, 2010; Levi, 2008; Petrov, Popple, & McKee, 2007; Petrov & Meleshkevich, 2011a; Petrov & Meleshkevich, 2011b; Shechter & Yashar, 2021; but see Strasburger, 2020; Strasburger & Malania, 2013).

Proposed models for explaining inner–outer asymmetry include cortical magnification, receptive field size, and spatial attention. According to the cortical magnification view, crowding is due to smaller

critical distances in the periphery (Motter & Simoni, 2007; Pelli, 2008). Thus, inner–outer asymmetry reflects a smaller cortical distance between the outer flanker and the target compared with that between the inner flanker and the target. However, this view was challenged by the finding that inner–outer asymmetry is related to the reported item rather than the display. Namely, within the same display (i.e., the same cortical distance), crowding interference is substantially reduced when observers report the outer item rather than the central one (Shechter & Yashar, 2021). Moreover, the cortical magnification factor in V1 and other retinotopic areas is not large enough to explain the strong inner–outer asymmetry (Petrov et al., 2007).

Receptive field (RF) size accounts of asymmetry rely on the fact that RF size increases at the visual periphery (Chaney et al., 2014; Dayan & Solomon, 2010). For example, according to an optimal Bayesian model, the larger RF size of the outer flankers increases the number of RFs responding to the outer flanker compared to the inner one and thereby biases the Bayesian selection toward the outer flanker (Dayan & Solomon, 2010).

According to the attentional selection view, attention is biased outward, leading to stronger interference of the outer flanker (Petrov & Meleshkevich, 2011b). Support for this view comes from findings showing that asymmetry is reduced when attention is biased inward, either by a foveal task or by blocking stimulus eccentricity (Petrov & Meleshkevich, 2011a). However, these manipulations involve other processes besides attention, such as task demand, stimulus uncertainty, and difficulty. Moreover, in this study, as with most demonstrations of asymmetry, the target was flanked by a single flanker and such a crowded display may bias attention. So, it is still unclear whether and how the locus of covert spatial attention affects the inner–outer asymmetry in a typical crowded display where both flankers are presented simultaneously. In the present study, we address this issue by manipulating covert attention.

Researchers manipulate covert spatial attention by presenting a peripheral cue that appears at the target location (valid), a non-target location (invalid), or the fixation location (neutral). The effect of attention is assessed by comparing valid and neutral trials, and the cost of inattention is measured by comparing neutral and invalid trials. Behavioral studies showed that attention enhances spatial resolution (e.g., Yeshurun & Carrasco, 1998). Neurophysiological investigations suggest that this enhancement in spatial resolution can be linked to two possible causes: either attention shrinks the receptive field size of cells over the attended location or it shifts the receptive field profile of cells toward the attended location (for a review, see Anton-Erxleben & Carrasco, 2013).

Despite the compelling evidence for an attentional effect on spatial resolution, investigations of attentional manipulation in crowding have yielded mixed results. Some studies have shown that attention decreases critical spacing (i.e., the minimum spacing required for crowding) (Grubb, Behrmann, Egan, Minshew, Heeger, & Carrasco, 2013; Kewan-Khalayly, Migó, & Yashar, 2022; Yeshurun & Rashal, 2010); however, other studies have failed to demonstrate an effect on crowding (Scolari, Kohnen, Barton, & Awh, 2007; Strasburger, 2005). One possible reason for this inconsistency is that the locus of attention with respect to the target could have varied across these studies and therefore caused the attentional effect on crowding. Indeed, in Scolari et al. (2007) the cue appeared at the location of the target, whereas in Yeshurun and Rashal (2010) the cue appeared closer to the fovea (inner flanker) than the target.

The exact locus of attention also has implications for investigating the predictions of the two competing asymmetry models. The attentional account predicts a direct positive effect between attention and asymmetry—namely, directing covert attention toward the outer flanker will increase asymmetry, whereas directing covert attention toward the inner flanker will decrease asymmetry. By contrast, the RF size view predicts a reduction of asymmetry when the locus of attention is at the outer-flanker location due to a change in either the size or the profile of the RFs over the attended location.

Finally, the effect of locus of attention may vary across the different types of crowding errors. Investigations of the pattern of crowding errors have revealed that crowding often leads to the misreporting of a flanker as the target (Ester, Klee, & Awh, 2014; Freeman, Chakravarthi, & Pelli, 2012; Harrison & Bex, 2015; Jimenez, Kimchi, & Yashar, 2022; Strasburger & Malania, 2013; Yashar et al., 2019). However, the effect of crowding on the perception of orientation, color, spatial frequency (SF), and motion is distinctive (Greenwood & Parsons, 2020; Yashar et al., 2019). For example, with orientation observers often misreport a flanker as the target, whereas with SF they average flankers and target values together (Yashar et al., 2019). Importantly, crowding of two features, such as orientation and color, leads to misbinding errors, such as reporting an orientation of one item with the color of another (Yashar et al., 2019). The binding process is particularly relevant for the issue of attention, as spatial attention is considered to play a critical role in feature binding. Attention is thought to act as a glue that binds features together (Treisman & Gelade, 1980; Treisman & Schmidt, 1982), perhaps by increasing spatial resolution through the reduction of RF size (Reynolds & Desimone, 1999).Thus, understanding the role of attention in crowding may shed light on its role in the feature binding process. However, whether

and how the locus of attention affects misreport and binding error in crowding is still unclear.

In the present study, we tested whether and how the locus of covert spatial attention affects feature binding in crowded displays and how it interacts with inner–outer asymmetry. In two experiments, observers estimated either the orientation of a grating (Gabor) stimulus (Experiment 1) or the tilt and color of a T-shaped object (Experiment 2) by reporting each feature in a continuous space. The target appeared on the horizontal meridian at 7° eccentricity and was either alone (uncrowded) or along with two flankers, one on each side of the target on the horizontal meridian (a radial crowding). We used a center-to-center distance between the target and the flankers—that is, within the crowding window (I, e, $m < 0.5$ of eccentricity). To manipulate attention, we presented a pre-cue (a circle) at one of four possible locations: fixation (neutral cue), the inner-flanker location (inner cue), the target location (target cue), and the outer-flanker location (outer cue). To assess the pattern of crowding and binding errors, we compared the fitting of probabilistic mixture modeling to the error distributions. The results reveal that binding errors in a radial crowded display reflect the inner–outer asymmetry. Analysis of cue positions revealed that the effect of covert attention is contingent on the locus of attention within the crowded stimulus and that covert attention is involved in the inner–outer asymmetry. The methods used and the data analyzed in the present study are available in the Open Science Framework repository (https://osf.io/ck4b2/?view_only=ce9737fce46046238b273fd4d66c71ad).

Experiment 1

Method

Observers

Eighteen students (6 males; age range 22–35 years, $M = 26.66$, $SD = 3.49$) from the University of Haifa participated in this experiment for either course credit or payment of 50 shekels (around \$14) per hour. We estimated that a sample size of 16 observers was required to detect a medium to large effect with 80% power, given a 0.05 significance criterion based on a priori power analysis using effect sizes from previous studies (Shechter & Yashar, 2021; Yashar et al., 2019). We collected data from two more observers because of possible dropouts or equipment failure. All observers were naïve to the research question and reported normal or corrected-to-normal vision, with no reported attention deficits. We obtained written informed consent from all observers before the experiment. The University Committee on Activities Involving Human Subjects at Haifa University approved all experimental procedures.

Apparatus

We used MATLAB (MathWorks, Inc., Natick, MA) with the Psychophysics Toolbox extensions (Kleiner, Brainard, Pelli, Ingling, Murray, & Broussard, 2007) to generate the stimuli and task. We ran the experiment on an iMac (Apple, Cupertino, CA) connected to a gamma-corrected 21-inch CRT monitor (with 1280 × 960 resolution and 85-Hz refresh rate). We used the EyeLink 1000 (SR Research, Kanata, ON, Canada), an infrared eye tracker, to monitor and record eye movement and a SpectroCAL MKII (Cambridge Research Systems, Rochester, UK) spectroradiometer to calibrate luminance and color. Observers were tested individually in a dimly lit room and used a mouse to generate responses. We used a chinrest to set the viewing distance of each observer at 57 cm.

Stimuli and procedure

Figure 1 illustrates the sequence of events in a trial. All stimuli were presented on a gray background (luminance 53 cd/m^2). Each trial began with the fixation display, a black dot (subtending 0.24° of visual angle, luminance 0.0073 cd/m^2) at the center of the screen. Following observer fixation for 300 ms, a pre-cue (a black circle 1.8° in diameter and 0.5° pen width) appeared for 50 ms. The cue appeared at the location of the upcoming target (target cue 25% of the trials) or the less eccentric flanker (inner cue 25% of the trials) or the more eccentric flanker (outer cue 25% of the trials). In the remaining 25% of the trials, the cue appeared at fixation (neutral cue). After an interstimulus interval of 50 ms, a peripheral target appeared for 100 ms. The target was a Gabor: a sinusoidal grating (1.5 c/°) with a Gaussian envelope ($SD = 0.65°$) and 75% contrast (the size was about 1.8° in diameter). The target was located on the horizontal meridian with 7° of eccentricity in either the left or the right hemifield. The target appeared either alone (uncrowded-display condition) or flanked by two Gabors (crowded-display condition). The flankers appeared one on each side of the target on the horizontal meridian. The center-to-center spacing between the target and the flankers was 2.3°.

Target and flanker orientation were randomly sampled from a circular parameter space of 180 values evenly distributed between 1° and 180°, with the restriction that, in each trial, the orientation of the Gabors differed by at least 15° from each other. A blank interval of 500 ms followed the stimulus display, then the response display appeared and remained on the screen until the observers completed their response. The response display was comprised of a probe
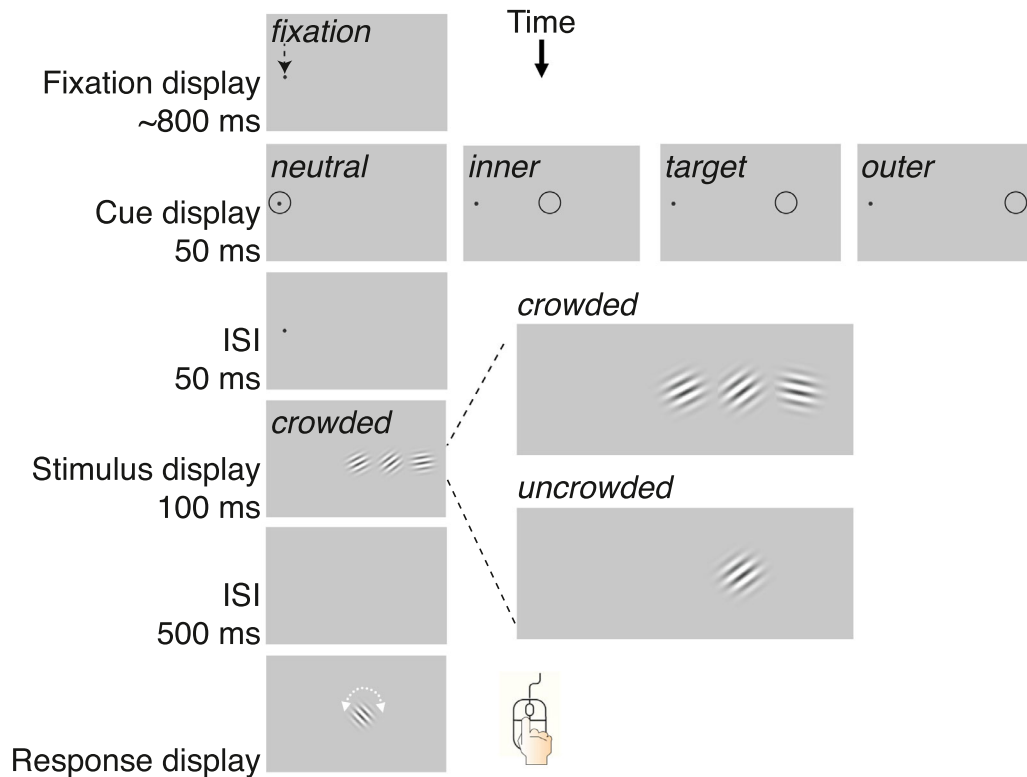
Figure 1. Illustration of the sequence of events within a trial in Experiment 1. Here, only the right hemifield of the display is shown (i.e., the fixation mark was at the center of the screen). The cue appeared at fixation (neutral), the inner-flanker location, the target location, or the outer-flanker location. The target (7° eccentricity) appeared alone (uncrowded) or radially flanked by two Gabors (2.5° center-to-center distance). Observers estimated the target orientation by adjusting the probe using a mouse.

(a Gabor at the center of the screen). Observers were asked to estimate the target orientation by adjusting the orientation of the probe using the mouse. Each condition had 100 trials (800 trials overall). Each observer completed 10 blocks of 80 trials in one session. In each block, there were 20 trials from each of the four cue conditions. The experiment began with a 40-trial practice block. Observers were encouraged to take a short break between blocks. To monitor eye fixation and stimulus eccentricity, we used online eye-tracking. We terminated trials in which the observer broke fixation and reran them at the end of the last block ($>2°$ from fixation).

Models and analysis

We calculated the estimation error in each trial by subtracting the true value of the target from the estimated value. First, for each observer in each condition, we assessed report bias and report precision by calculating the mean and the inverse of the standard deviation (std^{-1}) of the error, respectively. We then analyzed the error distributions by individually fitting probabilistic mixture models developed from the standard and standard-with-

misreporting models (Bays, Catalao, & Husain, 2009; Zhang & Luck, 2008).

For uncrowded-display trials we fitted the standard model (Equation 1), which uses a von Mises (circular) distribution to describe the probability density of the pooling estimation of the orientation of the target and a uniform component to reflect the guessing in estimation. The model has two free parameters (γ, σ). In this model, the probability of reporting a feature value $p(\hat{\theta})$ is

$$p\left(\hat{\theta}\right) = (1 - \gamma)\phi_\sigma\left(\hat{\theta} - \theta\right) + \gamma\left(\frac{1}{n}\right) \quad (1)$$

where $\hat{\theta}$ is the value of the reported feature and θ is the actual value of the target feature, γ is the proportion of trials in which observers are randomly guessing (guessing rate) with $n = 180$, ϕ_σ is the von Mises distribution with a standard deviation σ (variability) and a mean of 0. For crowded-display trials, we compared the fitting of models that included a component of misreporting a flanker as the target.

The one-misreport model (Equation 2) has three free parameters (γ, σ, β). This model adds a misreporting

component to the standard model. The misreport component describes the probability of reporting one of the flankers to be the target. In this model, the probability of reporting a feature value is

$$p\left(\hat{\theta}\right) = (1 - \gamma - \beta)\phi_\sigma\left(\hat{\theta} - \theta\right)$$
$$+ \gamma\left(\frac{1}{n}\right) + \frac{1}{m}\beta\sum_{i=1}^{m}\phi_\sigma\left(\hat{\theta} - \varphi_i\right) \quad (2)$$

where β is the probability of misreporting a flanker as the target, and φ_i is the actual value of the i flanker and m is the total number of flankers. The variability of the distribution around each stimulus was assumed to be the same.

The two-misreport model (Equation 3) has four free parameters (γ, σ, β^{In}, β^{Out}). The model adds two misreporting components to the standard model. Each misreport component describes the probability of reporting one of the flankers to be the target. In this model, the probability of reporting a feature value is

$$p\left(\hat{\theta}\right) = \left(1 - \gamma - \beta^{\ln} - \beta^{out}\right)\phi_\sigma\left(\hat{\theta} - \theta\right)$$
$$+ \gamma\left(\frac{1}{n}\right) + \beta^{In}\phi_\sigma\left(\hat{\theta} - \varphi^{In}\right)$$
$$+ \beta^{out}\phi_\sigma\left(\hat{\theta} - \varphi^{out}\right) \quad (3)$$

where β^{In} is the probability of misreporting the inner flanker as the target, and β^{Out} is the probability of misreporting the outer flanker as the target, φ^{In} and φ^{Out} are the actual value of the inner flanker and the outer flanker, respectively.

In addition to the standard model, the one-misreport model, and the two-misreport model, we also fitted these models with a target bias component: the bias standard model, the bias one-misreport model, and the bias two-misreport model. These models were similar to the regular models except that the mean (μ) of the von Mises distribution around the target ($\phi_{\sigma,\mu}$) was a free parameter.

The standard with bias model (Equation 4) has three free parameters (μ, γ, σ):

$$p\left(\hat{\theta}\right) = (1 - \gamma)\phi_{\sigma,\mu}\left(\hat{\theta} - \theta\right) + \gamma\left(\frac{1}{n}\right) \quad (4)$$

The one-misreport with bias model (Equation 5) has four free parameters (μ, γ, σ, β):

$$p\left(\hat{\theta}\right) = (1 - \gamma - \beta)\phi_{\sigma,\mu}\left(\hat{\theta} - \theta\right)$$
$$+ \gamma\left(\frac{1}{n}\right) + \frac{1}{m}\beta\sum_{i=1}^{m}\phi_\sigma\left(\hat{\theta} - \varphi_i\right) \quad (5)$$

The two-misreport with bias model (Equation 6) has five free parameters (μ, γ, σ, β^{In}, β^{Out}):

$$p\left(\hat{\theta}\right) = \left(1 - \gamma - \beta^{In} - \beta^{Out}\right)\phi_{\sigma,\mu}\left(\hat{\theta} - \theta\right)$$
$$+ \gamma\left(\frac{1}{n}\right) + \beta^{In}\phi_\sigma\left(\hat{\theta} - \varphi^{In}\right)$$
$$+ \beta^{Out}\phi_\sigma\left(\hat{\theta} - \varphi^{Out}\right) \quad (6)$$

We used the MemToolbox (Suchow, Brady, Fougnie, & Alvarez, 2013) for model fitting and comparison. To compare models, we calculated Akaike information criterion with correction (AICc) for the individual fits. We calculated the target reporting rate as $Pt = (1 - \gamma)$, $Pt = (1 - \gamma - \beta)$, and $Pt = (1 - \gamma - \beta^{In} - \beta^{Out})$ in the standard model, one-misreport model, and two-misreport model, respectively.

We performed a 2×4 analysis of variance (ANOVA) with display condition (uncrowded display vs. crowded display) and cue condition (neutral, inner, valid, outer) as within-subject factors on precision and parameters of the best-performing model. To test our main hypothesis, we performed a four-way ANOVA with cue position as the within-subject factor on the best-fitting model parameters.

Results and discussion

Figure 2A plots the distribution of errors for the uncrowded- and crowded-display trials. The mean report bias in all conditions was within the range of $\pm 2°$, indicating that there was no systematic report bias (Supplementary Table S1). First, we performed a 2×4 repeated-measures ANOVA with display condition and cue position as within-subject factors on precision. As expected, report precision was higher in the uncrowded-display trials than in the crowded-display trials (Figure 2B). There were no significant main effect of cue position and no significant interaction between cue position and display condition on precision (all $p > 0.72$) (Supplementary Table S1). Planned one-way ANOVA with cue position on precision showed no main effect of cue position in uncrowded display trials ($F < 1$) (Figure 2C), whereas cue position significantly affected precision in crowded trials, $F(3, 45) = 5.49$, $p = 0.002$, $\eta^2 = 0.3$, with the highest precision in inner-cue trials and lowest precision in outer-cue trials (Figure 2D).

Model fitting results

For crowded-display trials, as indicated by the mean AICc (Figure 3A), the two-misreport model outperformed the standard model and the one-misreport models, suggesting that the misreport rate differed between the inner and the outer flankers (Figure 3A). Next, we analyzed model parameters (Figures 2D and 2E, Figures 3B and 3C, Supplementary Table S2) of the standard model in uncrowded-display trials and the two-misreport model in crowded-display
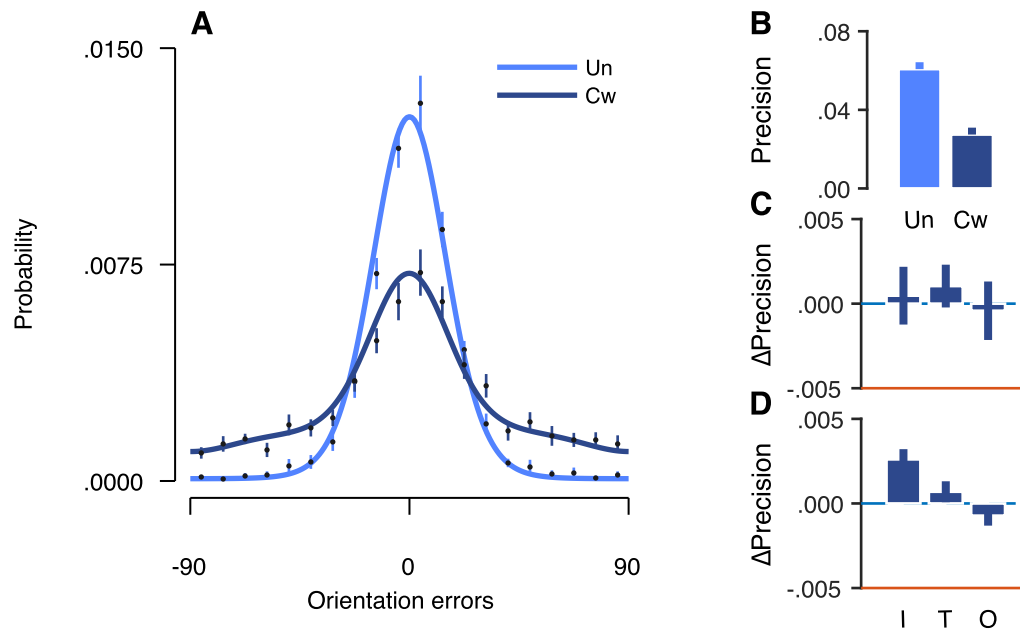
Figure 2. Uncrowded-display versus crowded-display trials in Experiment 1. (A) Merror distribution of errors (dark dots) in the uncrowded-display trials and the crowded-display trials. Errors were binned into 20 equal-width (9°) bins. Solid lines are the standard model and the two-misreport model for uncrowded and crowded trials, respectively. Model lines were generated best on the average parameters of the individual fits. (B) Precision (std^{-1} of report errors in degrees) as a function of display condition. Un = uncrowded-display trials, Cw = crowded-display trials. (C, D) Cueing effect (peripheral cue − neutral cue) for uncrowded (C) and crowded (D) trials. Error bars, ±1 within-subject *SE*.
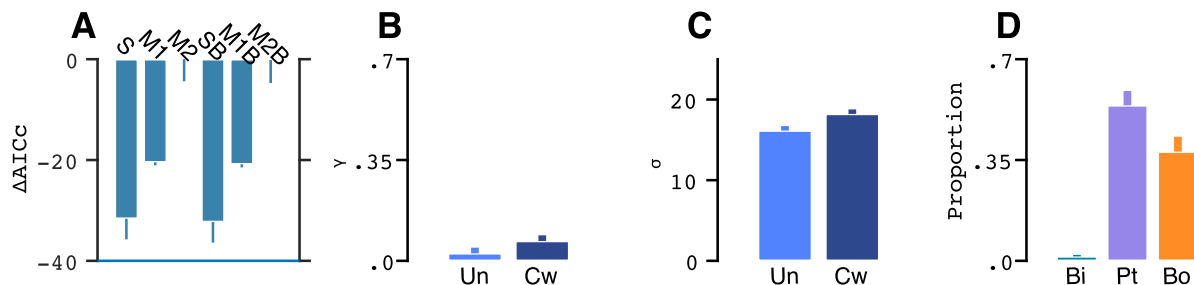

Figure 3. Model comparisons and parameters in Experiment 1. (A) Model comparisons in crowded trials. △AICc was calculated by subtracting from each AICc the AICc of the best-performing model (two-misreport model). Lower △AICc values indicate better performance. S = standard, 2M = two-misreport, 1M = one-misreport, SB = standard with bias, 2M = two-misreport with bias, 1M = one-misreport with bias. (B, C) Mean fitted guess rate (g) for each crowding display condition (B) and variability (s) for each crowding display condition (C). Parameters were fitted individually with the best-performing models (standard in uncrowded trials and two-misreport in crowded trials). (D) Mean report component of the two-misreport model in crowded display trials. Bi = inner flanker, Pt = target, Bo = outer flanker. Error bars, ±1 within-subject *SE*.

trials. Crowded display increased variability (σ), $F(1, 15) = 14.99$, $p = 0.002$, $\eta^2_p = 0.5$ (Figure 3B), and guesses (γ), $F(1, 15) = 6.97$, $p = 0.019$, $\eta^2_p = 0.32$, and decreased target reports (*Pt*), $F(1, 15) = 113.93$, $p < 0.001$, $\eta^2_p = 0.88$ (Figure 3C). (See the Supplementary Materials for the remaining 2 × 4 ANOVA results on all parameters.) Figure 3D shows the misreport component in the crowded trials. The misreport rate

of the outer flanker was substantially higher than that of the inner flanker, demonstrating inner–outer asymmetry.

Next, we analyzed the cuing effect in crowded-display trials. Cue position modulated misreport of the outer flanker (β^{Out}), $F(3, 45) = 8.42$, $p < 0.001$, $\eta^2_p = 0.56$, with higher β^{Out} in outer cue trials compared with inner and target cue trials (Figure 4A). There was no
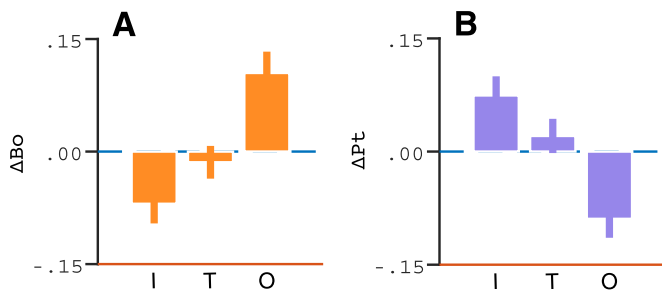
Figure 4. Cueing effects in crowded-display trials of Experiment 1. (A) Cueing effect on target report rate. (B) Cueing effect on the rate of misreporting the outer-flanker as target. Cueing effect was calculated by subtracting the report rate in the neutral cue from the report rate in each peripheral cue position. I = inner, O = outer. Error bars, ±1 within-subject *SE*.

significant effect on misreport of the inner flanker (β^{In}) ($F < 1$). There was a significant effect of cue position on target report rate (Pt), $F(3, 45) = 7.66$, $p < 0.001$,

$\eta^2_p = 0.51$ (Figure 4B). These findings show that the chance of reporting the outer flanker increased with the proximity of the outer flanker to the locus of focal attention, demonstrating a positive relation between attention and inner–outer asymmetry.

Experiment 2

In Experiment 2, we extended our investigation to the process of feature binding. A recent study (Yashar et al., 2019) tested the effect of crowding on feature binding—that is, the integration of feature dimensions (e.g., tilt, color, SF) to a coherent object. Observers performed a double report task by estimating both the tilt (fully circular space 0°–360°) and the color (on a color wheel 0°–360°) of a T-shaped target. In a crowded display, observers misreported tilt or colors in an independent manner; that is, observers often reported
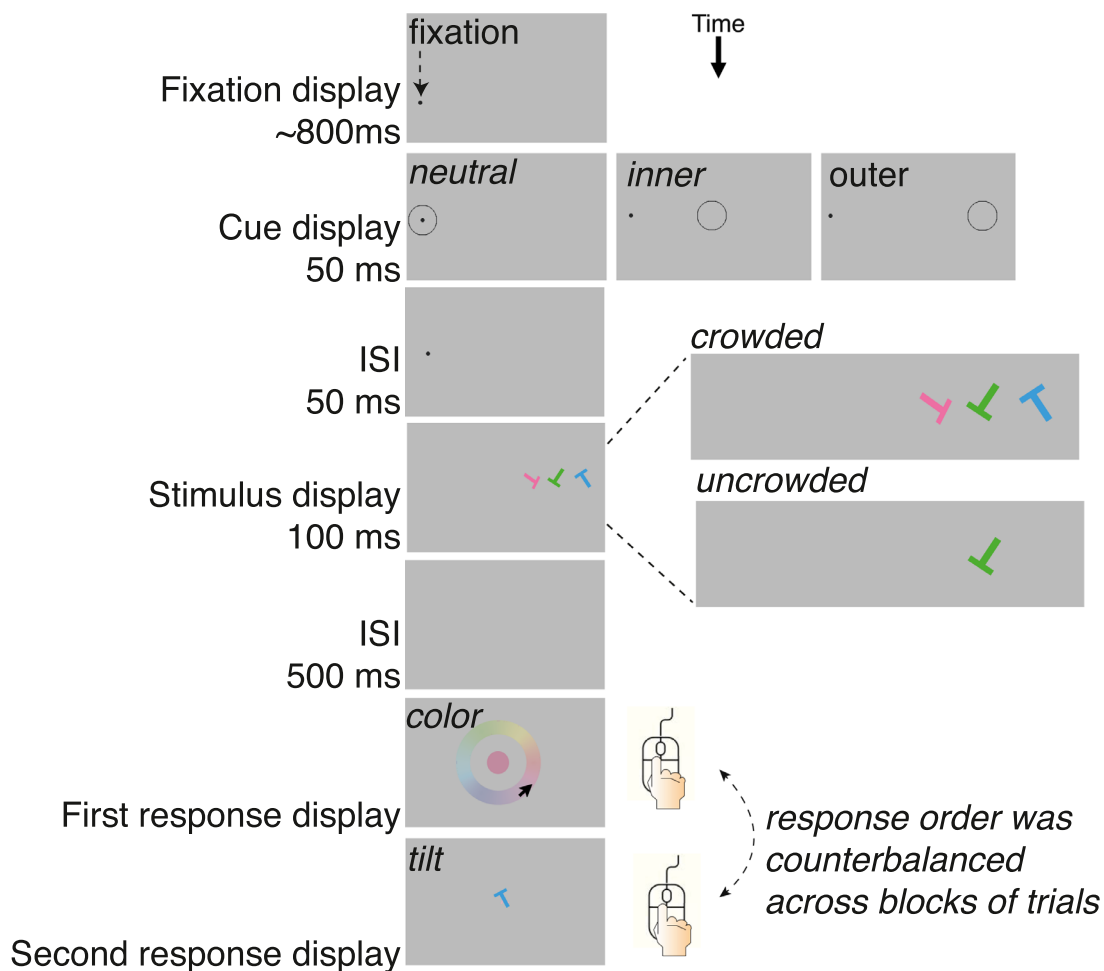

Figure 5. Illustration of the sequence of events within a trial in Experiment 2. The fixation point, here on the left, was presented at the center of the screen. The cue appeared at fixation (neutral), at the inner-flanker location, or at the outer-flanker location. Observers estimated the target color and orientation by adjusting the probe using a mouse.

presented features accurately but in an inaccurate conjunction. These findings suggest that crowding disrupts the integration of features into a coherent object, leading to what is known as misbinding or "illusory conjunction" (Treisman & Schmidt, 1982).

Classical visual attention views argue that spatial attention plays a critical role in feature binding. Indeed, when covert attention was disrupted, misbinding errors were reported in uncrowded displays (i.e., object spacing was outside the crowding window) (Treisman & Gelade, 1980; Treisman & Schmidt, 1982). In Experiment 2, we tested the effect of covert spatial attention on both crowding and binding errors. If binding errors in crowding are due to diffused attention, then directing spatial attention should modulate binding errors.

Method

The method was the same as in Experiment 1 except for the following.

Observers

Nineteen students (7 males; age range 23–37 years, $M = 28.37$, $SD = 4.34$).

Stimuli and procedure

Figure 5 illustrates the sequence of events in a trial. The pre-cue appeared at one of three possible locations: fixation (neutral), inner flanker, or outer flanker. The target and the flankers were T-shaped items, each subtending $1.8° \times 1.8°$ and drawn with a 0.3° stroke. The tilt and color of the target and the flankers were each independently selected at random from two circular parameter spaces. Target and flanker tilt were randomly selected out of 360 values evenly distributed between 1° and 360°. The color was randomly selected out of 360 values evenly distributed along a circle in the Derrington–Krauskopf–Lennie (DKL) color space (Derrington, Krauskopft, Lenniet, Kra, & Lennie, 1984). We followed the same color space calibration as Yashar et al. (2019) (see supplementary information in Yashar et al., 2019). Stimuli color and background were equiluminant.

The color response displays included a color wheel (2° thick with an inner radius of 5°) containing 360 colors. Observers were asked to estimate the target color by selecting a color on the color wheel using the mouse curser. During response, a visual presentation of the selected color was presented at the center of the screen. As in Experiment 1, observers estimated the target tile by rotating a T-shaped item at the center of the screen using the mouse.

In both response types, a final report was made by clicking the mouse. The response order (tilt and color) alternated every block of 150 trials. In each of the three cue positions there were 200 crowded-display trials and 100 uncrowded-display trials (900 trials in total). We encouraged observers to take a short break every 50 trials. The experiment began with a 40-trial practice block.

Models and analysis

To analyze each feature space separately, we performed the same model fitting procedure as described in Experiment 1. Table 1 shows all free parameters of the joint-distribution models. To analyze

Component category		Tilt	Color	Mixture component	Probability density function
Bound target	1	Target	Target	$T_t T_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \theta_t)\phi_{\sigma_c}(\hat{\theta}_c - \theta_c)$
Feature error	2	Uniform	Uniform	$U_t U_c$	$(\frac{1}{360})^2$
	3	Uniform	Target	$U_t T_c$	$\frac{1}{360}\phi_{\sigma_c}(\hat{\theta}_c - \theta_c)$
	4	Target	Uniform	$T_t U_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \theta_t)\frac{1}{360}$
	5	Outer	Uniform	$O_t U_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{Out})\frac{1}{360}$
	6	Inner	Uniform	$I_t U_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{In})\frac{1}{360}$
	7	Uniform	Outer	$U_t O_c$	$\frac{1}{360}\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{Out})$
	8	Uniform	Inner	$U_t I_c$	$\frac{1}{360}\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{In})$
Binding error	9	Outer	Inner	$O_t I_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{Out})\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{In})$
	10	Inner	Outer	$I_t O_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{In})\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{Out})$
	11	Outer	Target	$O_t T_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{Out})\phi_{\sigma_c}(\hat{\theta}_c - \theta_c)$
	12	Inner	Target	$I_t T_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{In})\phi_{\sigma_c}(\hat{\theta}_c - \theta_c)$
	13	Target	Outer	$T_t O_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \theta_t)\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{Out})$
	14	Target	Inner	$T_t I_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \theta_t)\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{In})$
Object error	15	Inner	Inner	$I_t I_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{In})\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{In})$
	16	Outer	Outer	$O_t O_c$	$\phi_{\sigma_t}(\hat{\theta}_t - \varphi_t^{Out})\phi_{\sigma_c}(\hat{\theta}_c - \varphi_c^{Out})$

Table 1. Mixture components of the joint-standard model (rows 1–4) and joint-misreport model (rows 1–16).

binding errors, we fitted joint mixture models (Bays, Wu, & Husain, 2011; Dowd & Golomb, 2019) to the joint distributions of tilt and color. For uncrowded-display trials, each feature dimension report could come from uniform or target distributions, leading to four possible report combinations of tilt and color distribution (Table 1, rows 1–4, joint-standard model). For crowded-display trials, each feature dimension report could come from one of four distributions: uniform, Gaussian over the target, Gaussian over the inner flanker or Gaussian over the outer flanker. Because we had two feature dimensions, the total number of possible distribution combinations of tilt and color was 16 (Table 1, rows 1–16, joint-misreport model). Each joint model also included a von Mises variability component for each feature dimension (σ_t, σ_c). (Note that, because the sum of all report components is equal to 1, $T_t T_c$ (Table 1,

row 1) was defined as $T_t T_c = 1 - \sum_{i=2}^{16}(p_i)$, where p is the report probability of the ith component (Table 1, rows 2–16). Thus, overall, the joint-standard model had five free parameters, and the joint-misreport model had 17 free parameters.

We used the MCMC function in the MemFit toolbox to individually fit the models in each crowding and cue condition. To simplify the analysis of the joint-misreport model (16 components) in crowded-display trials, we grouped components into four categories of reports: (a) bound target (Table 1, row 1), which related to the report of both tilt and color of the target; (b) feature error (Table 1, rows 2–8), which related to any guessing component; (c) binding errors (Table 1, rows 9–14), which related to misreport of different items (e.g., the target tilt with a flanker color); and (d) object error (Table 1, rows 15 and 16), which
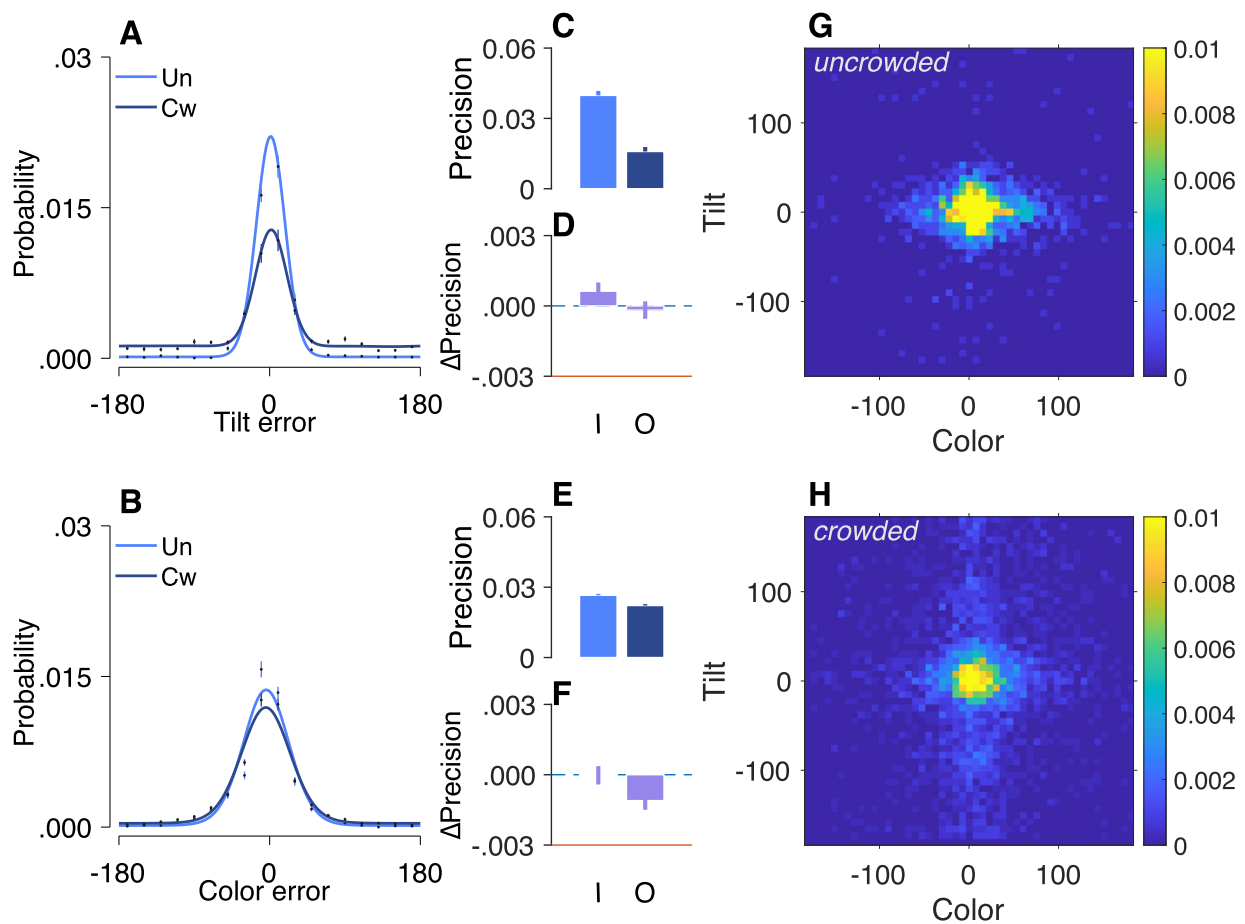

Figure 6. Uncrowded-display trials versus crowded-display trials in Experiment 2. (A, B) Mean error distributions (dark dots) in each crowding condition were plotted for tilt (A) and color (B) reports. Errors were binned into 20 equal-width (9°) bins. Solid lines plot the best fitted model in each condition and report feature. (C) Mean precisions, calculated as the inverse of the standard deviation of the report errors ($SD°^{-1}$) in uncrowded- and crowded-display trials for tilt. I = inner, O = outer. (D) Mean cueing effect on precision in crowded-display trials for tilt. (E) Mean precisions ($SD°^{-1}$) in uncrowded- and crowded-display trials for color. (F) Mean cueing effect on precision in crowded-display trials for color. We calculated the cueing effect by subtracting the neutral cue from each peripheral cue condition. (G) Joint distribution of tilt and color reports for uncrowded-display trials. (H) Joint distribution of tilt and color reports for crowded-display trials. Error bars, ±1 within-subject *SEM*.

related to misreporting both features of the same flanker. Note that both object error and target report reflect correct binding. To test for the effect of covert spatial attention on binding, we analyzed the effect of cue position on each of the four error types. We performed a three-way, repeated-measure ANOVA with cue position as a within-subject factor on each component category. To test for the effect of cue position on correct binding in uncrowded display trials, we performed a three-way repeated measures ANOVA with cue position as the within-subject factor on the bound target rate (Table 1, row 1) of the standard joint model.

Results and discussion

Figures 6A and 6B plot the distribution of errors for the uncrowded- and crowded-display trials for tilt and color. For either tilt or color reports, precision was higher in the uncrowded-display trials than in the crowded-display trials, $F(1, 18) = 78.75$, $p < 0.001$, $\eta^2_p = 0.81$ and $F(1, 18) = 24.98$, $p < 0.001$, $\eta^2_p = 0.58$, respectively (Figures 6C and 6E). The main effect of cue position on precision was significant for color, $F(2, 36) = 5.15$, $p < 0.019$, $\eta^2_p = 0.22$ (Figures 6D

and 6F). No other effect was significant on precision, and no effect was significant on report bias, (all $p > 0.1$) (Supplementary Table S3).

Model fitting results

For tilt report in crowded-display trials, as indicated by the mean AICc (Figure 7A), the two-misreport model with bias outperformed models with one or no misreport component, suggesting that, as for Gabor orientation in Experiment 1, the misreport rate for tilt differed between the inner and the outer flankers. For color, the standard with bias model outperformed the other models. We removed one observer from the model fitting analysis due to a high guessing rate (>0.50). Thus, for tilt, we analyzed the fitted parameters of the standard with bias model and the bias two-misreport models in uncrowded- and crowded-display trials, respectively. For color, we analyzed the fitted parameters of the standard with bias model (see all fitted parameters values in Supplementary Tables S4 and S5).

Crowding increased the tilt guess rate (γ) and variability (σ) (Figures 7B and 7C), $F(1, 17) = 53.98$, $p < 0.001$, $\eta^2_p = 0.76$ and $F(1, 17) = 4.58$, $p = 0.047$, $\eta^2_p = 0.21$, respectively. All other effects on tilt σ were not
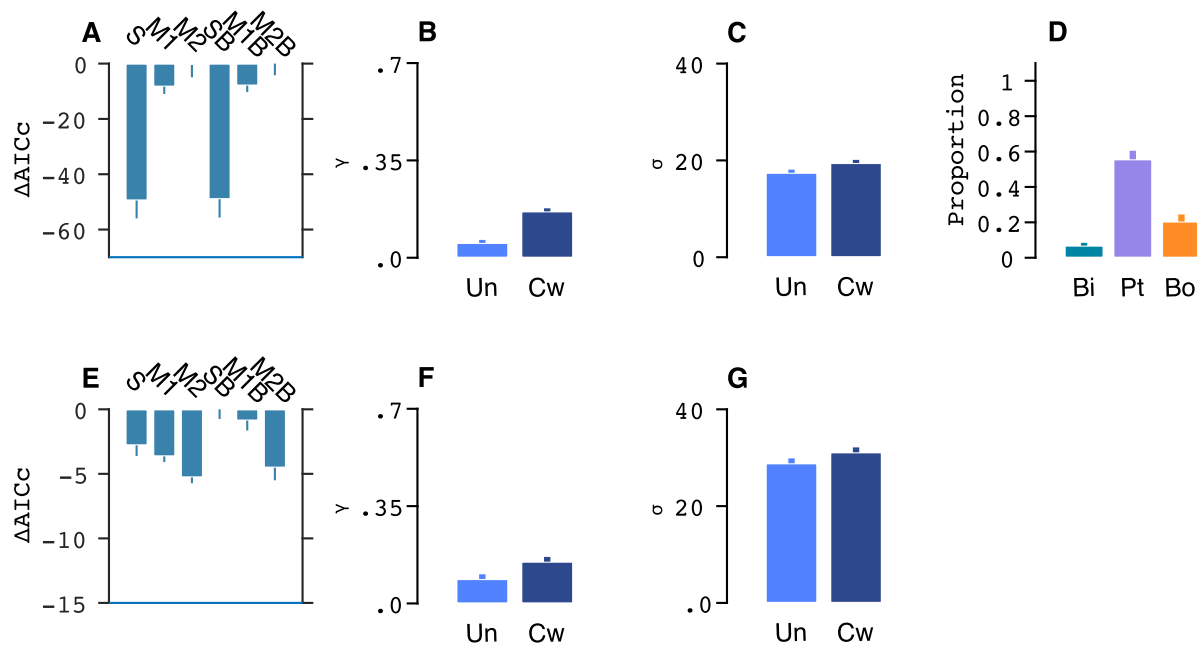

Figure 7. Model comparisons and parameters in Experiment 2. (A) Model comparisons in crowded trials for tilt. For each model, △AICc was calculated by subtracting the AICc of the best-performing model (two-misreport model). Lower △AICc indicates better performance. S = standard, 2M = two-misreport, 1M = one-misreport, SB = standard with bias, 2M = two-misreport with bias, 1M = one-misreport with bias. (B) Mean fitted tilt guess rate (γ). (C) Mean fitted tilt variability (σ). (D) Mean tilt report component of the two-misreport model in crowded display trials. Bi = inner flanker, Pt = target, Bo = outer flanker. (E) Model comparisons in crowded trials for color. For tilt and color, parameters were fitted individually with the best-performing models (standard model for tilt in uncrowded trials and for color in uncrowded and crowded trials, and two-misreport model for tilt in crowded trials). Error bars, ±1 within-subject SE.
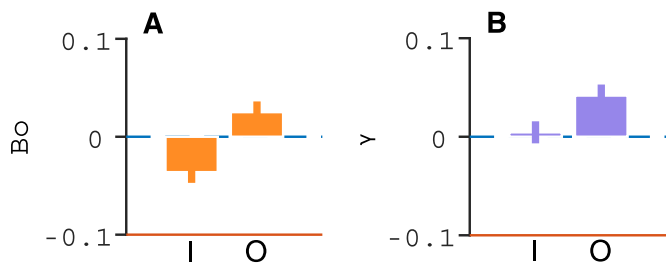
Figure 8. Cueing effects in crowded-display trials of Experiment 2. (A) Cueing effect on the rate of misreporting the outer-flanker tilt as target tilt. (B) Cueing effect on target-color guess rate. Cueing effect was calculated by subtracting the report rate in the neutral cue from the report rate in each peripheral cue position. I = inner, and O = outer. Error bars, ±1 within-subject *SE*.

significant ($p > 0.1$). Next, we analyzed the cueing effect in crowded-display trials. Figure 8A plots β^{In}, Pt, and β^{Out} in neutral cue trials in crowded-display trials. Cue position significantly affected β^{Out}, $F(2, 34) = 9.07$, $p < 0.001$, with higher β^{In} in the outer cue trials than the inner cue and neutral cue trials. These findings indicate that the outer cue increases the inner–outer asymmetry. The effect of cue position on Pt (Figure 6B) and β^{In} was not significant, $F(2, 34) = 2.93$, $p = 0.065$, and $F(2, 34) = 2.94$, $p = 0.066$, respectively.

The effect of crowding on color was reflected in a lower guess rate (γ) in the uncrowded trials than in the crowded trials, $F(1, 17) = 6.69$, $p = 0.017$, $\eta^2_p = 0.28$ (Figure 7F). There was a main effect of cue position on guess rate, $F(2, 34) = 5.76$, $p = 0.006$, $\eta^2_p = 0.25$, with a higher guess rate in the outer cue trials compared with the cue trials (Figure 8B). This finding suggests that, as

with tilt and orientation (Experiment 1), the outer cue increases the hindering effect of crowding on feature recognition. No other main effect or interaction was significant (all $p > 0.05$).

Joint-distribution models

Figure 9A summarizes the joint report components based on the type of report error. A large proportion of the errors reflect binding errors (i.e., reporting the orientation of one item with the color of another item). These errors are mainly driven by misreports of the orientation of the outer flanker while reporting the color of the target (see Supplementary Table S7), indicating that the inner–outer asymmetry reflects feature binding errors in radial crowding.

Next, we tested the effect of cue position on binding error by analyzing the fitted parameters of the joint-distribution models (Supplementary Table S6). For the joint-standard model in uncrowded-display trials, there was no significant effect of cue position on the bound target report rate ($F < 1$). Figures 6G and 6H depict the joint distribution of tilt and color for uncrowded- and crowded-display trials, respectively.

Figure 9A shows the mean rate for each mixture component category in neutral trials with a crowded display. The bound target report rate was higher in the inner-cue position compared to the outer-cue position, $t(17) = 2.21$, $p = 0.04$ (Figure 9B). These findings reflect the overall increase in target report rate in inner-cue trials within each feature space. Interestingly, when cue position was tested on each of the three error types, cue position did not significantly affect the feature errors, binding errors, or object errors (all $p > 0.18$) (Figure 9B).
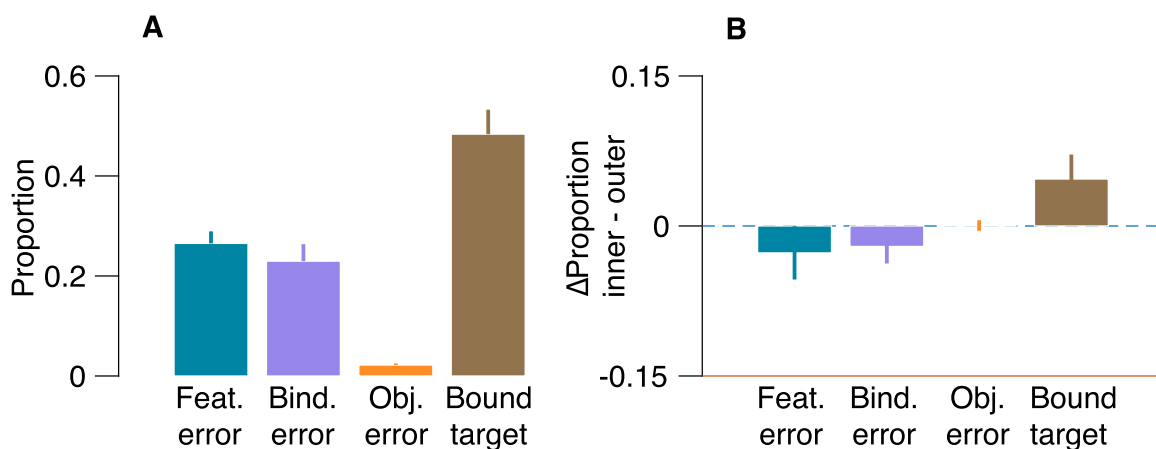

Figure 9. Joint feature report rates in Experiment 2. (A) Mean report rates for each report component category of the joint-misreport model. (B) For each component, we plotted the cueing effect by subtracting the report proportion in the outer-cue trials from the inner-cue trials. Feat. error = feature error, Bind. error = binding error, Obj. error = object error, Bound target = target reported in both features. Error bars, ±1 within-subject *SE*.

General discussion

The results show that the locus of covert attention within the crowded stimulus (i.e., inner, target, or outer locations) determines target identification. First, as in the study by Shechter and Yashar (2021), instead of the target, observers often misreported the orientation or tilt of the outer flanker rather than the inner one, which demonstrates inner–outer asymmetry in a typical display of radial crowding. Second, as in the study by Yashar et al. (2019), color crowding was substantially smaller than orientation crowding. Moreover, color errors did not reflect inner–outer asymmetry. Third, compared to maintaining allocation at fixation, attending to the target location did not affect target performance. Interestingly, attending to the inner-flanker location—a more foveal location than the target—increased the target identification and reduced the asymmetry. Finally, attending to the outer flanker reduced the target identification and increased the asymmetry.

This direct and positive relationship between covert attention and inner–outer asymmetry is consistent with the attentional bias account (Petrov & Meleshkevich, 2011a) and inconsistent with the receptive field size view of inner–outer asymmetry (Chaney et al., 2014; Dayan & Solomon, 2010). Whereas attentional accounts predict an increase in asymmetry when the locus of attention is the outer-flanker location, the receptive field size view predicts a decrease in asymmetry in outer-cue trials due to the reduction in receptive field size over the outer flanker.

Spatial attention and crowding

The present findings explain the inconsistent results of spatial attention investigations in crowding and attribute them to variations in the locus of attention. Consistent with the present study, previous studies using various cue and target types showed that cueing attention at the target location did not reduce crowding interference (Scolari et al., 2007; Strasburger & Malania, 2013), whereas cueing attention at location inner to the target reduced crowding interference (Grubb et al., 2013; Kewan-Khalayly et al., 2022; Yeshurun & Rashal, 2010).

The results of the present study cannot be explained by forward masking created by the cue. First, we used an empty circle shape for the cue, which was shown to be ineffective for forward masking of an orientation grating stimulus (Saarela & Herzog, 2008). Second, the pattern of results is inconsistent with a forward masking effect. In particular, when the cue appeared at the outer flanker location, observers more frequently misreported the outer flanker as the target compared to trials in which the cue appeared away from the outer flanker. By contrast, a masking effect by the cue predicted the reduction misreport rate of each cued item, and it is difficult to explain why the cue masked the target but not the outer flanker. The idea that attention is biased outward, therefore, provides a more parsimonious explanation and is consistent with studies of the inner–outer asymmetry phenomenon (Petrov & Meleshkevich, 2011a).

Investigations of the effect of spatial attention on basic signal processing typically display a simple stimulus around threshold levels, often by reducing stimulus strength (for a review, see Carrasco, 2011). Here, to test the attentional effect on crowding alone, we used a high-contrast super-threshold target. Thus, we did not expect to find an attentional effect in the uncrowded-display trials in which performance was relatively at ceiling levels (i.e., *Pt* close to 1).

Crowding, binding, and attention

The present study has implications for the feature-binding issue. First, this study replicates the results of Yashar et al. (2019) by showing that crowding errors reflect binding errors. Specifically, here, when observers had to report the tilt and color of a T-shaped target, they often misreported the tilt of the outer flanker with the color of the target (i.e., reporting an "illusory conjunction," a binding error). A joint misreport mixture model revealed that observers performed binding errors (i.e., reporting one feature from one object and the second feature from another object) or feature errors (i.e., reporting one or two feature values unrelated to a presented object—a guess). Notably, only a small percentage of trials reflected object error (i.e., reporting two features of the same flanker), suggesting that misreport errors reflect feature integration errors rather than confusion between coherent objects. This finding suggests that crowding is due to excessive integration processes and supports pooling models (Freeman et al., 2012; Freeman & Simoncelli, 2011; Keshvari & Rosenholtz, 2016; Rosenholtz, Yu, & Keshvari, 2019).

Second, the results provide insight into the role of spatial attention in feature binding. A prominent view of feature binding and attention is the feature integration theory (Treisman & Gelade, 1980), according to which attention operates as the "glue" that binds features together. Thus, this view predicts that allocating covert attention toward the crowded stimulus will reduce binding errors—namely, attention would lead to higher reports of a coherent target (bound target) and coherent flankers (object errors). However,

the present study results are inconsistent with this prediction. First, they showed that binding errors in a crowded display were persistent even when covert attention was allocated toward the crowded stimulus. Second, they showed that cueing covert attention to the crowded stimuli affected feature and binding errors but not object errors. In other words, when observers misreported a flanker feature as the target, covert spatial attention did not "glue" the other feature dimension to generate a coherent object perception and reports of the other feature of the same flanker (object errors).

This finding is inconsistent with previous studies that tested the effect of spatial attention on binding errors and showed that cueing attention decreased binding errors in a conjunction detection task (Briand, 1998; Prinzmetal, Presti, & Posner, 1986). However, these studies did not monitor eye movement and used a cue–stimulus stimulus onset asynchrony of 227 to 250 ms, which was enough time for a saccade (e.g., Mayfrank, Kimmig, & Fischer, 1987). Thus, it is unclear whether the cueing effect is due to overt (rather than covert) attention. Here, by monitoring eye movement using an eye tracker, we were able to test the effect of covert attention per se on feature binding. We showed that the effect of covert attention on feature binding is limited to the task-relevant item—the target.

Moreover, we showed both the cost and benefit of attentional allocation in a crowded display by testing various cue positions. Therefore, our feature binding findings go beyond a particular cue–target special relation. In Experiment 2, we used two cue positions, inner cue and outer cue, which we selected to maximize attentional cost and benefit based on Experiment 1. The pattern of cue position effect on crowding errors was consistent across experiments. Thus, it is unlikely that adding target–cue position in Experiment 2 would have changed the pattern of feature binding results.

Note that our findings mainly apply to bottom–up covert attention; however, top–down attention may still play a critical role in feature binding. Indeed, according to a prominent view, crowding is due to reduced attention resolution in the periphery (Chakravarthi & Cavanagh, 2007; He, Cavanagh, & Intriligator, 1996; Intriligator & Cavanagh, 2001; Tripathy & Cavanagh, 2002). This view assumes that the minimum size of the attentional selection region is larger in the periphery. Thus, when two or more items fall into the selection region, they are indistinguishable. Given that the selection region reflects top–down attention that differs from the bottom–up attention manipulated by the cue, our findings do not challenge the attentional selection view of crowding. Thus, reduced top–down attentional resolution may be responsible for binding errors in crowding.

Conclusions

The present study results reveal the important role of covert spatial attention in inner–outer asymmetry. The findings are consistent with the attentional bias account of inner–outer asymmetry and inconsistent with the receptive field size account. Our study also demonstrates a strong link between crowding and feature binding/integration and shows that crowding errors reflect binding errors.

Acknowledgments

References

Anton-Erxleben, K., & Carrasco, M. (2013). Attentional enhancement of spatial resolution: Linking behavioural and neurophysiological evidence. *Nature Reviews Neuroscience, 14*(3), 188–200, https://doi.org/10.1038/nrn3443.

Banks, W. P., Bachrach, K. M., & Larson, D. W. (1977). The asymmetry of lateral interference in visual letter identification. *Perception & Psychophysics, 22*(3), 232–240, https://doi.org/10.3758/BF03199684.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10):7, 1–11, https://doi.org/10.1167/9.10.7.

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

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177–178, https://doi.org/10.1038/226177a0.

Briand, K. A. (1998). Feature integration and spatial attention: More evidence of a dissociation between endogenous and exogenous orienting.

Journal of Experimental Psychology: Human Perception and Performance, 24(4), 1243–1256, https://doi.org/10.1037/0096-1523.24.4.1243.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*(13), 1484–1525, https://doi.org/10.1016/j.visres.2011.04.012.

Chakravarthi, R., & Cavanagh, P. (2007). Temporal properties of the polarity advantage effect in crowding. *Journal of Vision, 7*(2):11, 1–13, https://doi.org/10.1167/7.2.11.

Chaney, W., Fischer, J., & Whitney, D. (2014). The hierarchical sparse selection model of visual crowding. *Frontiers in Integrative Neuroscience, 8*, 73, https://doi.org/10.3389/fnint.2014.00073.

Chung, S. T. L. (2007). Learning to identify crowded letters: Does it improve reading speed? *Vision Research, 47*(25), 3150–3159, https://doi.org/10.1016/j.visres.2007.08.017.

Dayan, P., & Solomon, J. A. (2010). Selective Bayes: Attentional load and crowding. *Vision Research, 50*(22), 2248–2260, https://doi.org/10.1016/j.visres.2010.04.014.

Derrington, A. M., Krauskopft, J., Lenniet, P., Kra, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology, 357*, 241–265, https://doi.org/10.1113/jphysiol.1984.sp015499.

Dowd, E. W., & Golomb, J. D. (2019). Object-feature binding survives dynamic shifts of spatial attention. *Psychological Science, 30*(3), 343–361, https://doi.org/10.1177/0956797618818481.

Ester, E. F., Klee, D., & Awh, E. (2014). Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 1022–1033, https://doi.org/10.1037/a0035377.

Freeman, J., Chakravarthi, R., & Pelli, D. G. (2012). Substitution and pooling in crowding. *Attention, Perception, and Psychophysics, 74*(2), 379–396, https://doi.org/10.3758/s13414-011-0229-0.

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience, 14*(9), 1195–1201, https://doi.org/10.1038/nn.2889.

Gori, S., & Facoetti, A. (2015). How the visual aspects can be crucial in reading acquisition: The intriguing case of crowding and developmental dyslexia. *Journal of Vision, 15*(1):8, 1–20, https://doi.org/10.1167/15.1.8.

Greenwood, J. A., & Parsons, M. J. (2020). Dissociable effects of visual crowding on the perception of color and motion. *Proceedings of the National Academy of Sciences, USA, 117*(14), 8196–8202, https://doi.org/10.1073/pnas.1909011117.

Grubb, M. A., Behrmann, M., Egan, R., Minshew, N. J., Heeger, D. J., & Carrasco, M. (2013). Exogenous spatial attention: Evidence for intact functioning in adults with autism spectrum disorder. *Journal of Vision, 13*(14):9, 1–13, https://doi.org/10.1167/13.14.9.

Harrison, W. J., & Bex, P. J. (2015). A unifying model of orientation crowding in peripheral vision. *Current Biology, 25*(24), 3213–3219, https://doi.org/10.1016/j.cub.2015.10.052.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334–337, https://doi.org/10.1038/383334a0.

Hussain, Z., Webb, B. S., Astle, A. T., & McGraw, P. V. (2012). Perceptual learning reduces crowding in amblyopia and in the normal periphery. *Journal of Neuroscience, 32*(2), 474–480, https://doi.org/10.1523/JNEUROSCI.3845-11.2012.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*, 171–216, https://doi.org/10.1006/cogp.2001.0755.

Jimenez, M., Kimchi, R., & Yashar, A. (2022). Mixture-modeling approach reveals global and local processes in visual crowding. *Scientific Reports, 12*(1), 6726, https://doi.org/10.1038/s41598-022-10685-z.

Keshvari, S., & Rosenholtz, R. (2016). Pooling of continuous features provides a unifying account of crowding. *Journal of Vision, 16*(3):39, 1–15, https://doi.org/10.1167/16.3.39.

Kewan-Khalayly, B., Migó, M., & Yashar, A. (2022). Transient attention equally reduces visual crowding in radial and tangential axes. *Journal of Vision, 22*(9):3, 1–9, https://doi.org/10.1167/jov.22.9.3.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception, 36*(14), 1–16, https://doi.org/10.1068/v070821.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654, https://doi.org/10.1016/j.visres.2007.12.009.

Mayfrank, L., Kimmig, H., & Fischer, B. (1987). The role of attention in the preparation of visually guided saccadic eye movements in man. In J. K. O'Regan, & A. Levy-Schoen (Eds.), *Eye Movements from Physiology to Cognition* (pp. 37–45). Amsterdam: Elsevier.

Motter, B. C., & Simoni, D. A. (2007). The roles of cortical image separation and size in active visual search performance. *Journal of Vision, 7*(2):6, 1–15, https://doi.org/10.1167/7.2.6.

Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology, 18*(4), 445–451, https://doi.org/10.1016/j.conb.2008.09.008.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12):12, 1136–1169, https://doi.org/10.1167/4.12.12.

Petrov, Y., & Meleshkevich, O. (2011a). Locus of spatial attention determines inward-outward anisotropy in crowding. *Journal of Vision, 11*(4):1, 1–11, https://doi.org/10.1167/11.4.1.

Petrov, Y., & Meleshkevich, O. (2011b). Asymmetries and idiosyncratic hot spots in crowding. *Vision Research, 51*(10), 1117–1123, https://doi.org/10.1016/j.visres.2011.03.001.

Petrov, Y., Popple, A. V., & McKee, S. P. (2007). Crowding and surround suppression: Not to be confused. *Journal of Vision, 7*(2):12, 1–9, https://doi.org/10.1167/7.2.12.

Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration? *Journal of Experimental Psychology: Human Perception and Performance, 12*(3), 361–369, https://doi.org/10.1037/0096-1523.12.3.361.

Reynolds, J. H., & Desimone, R. (1999). The role of neural mechanisms of attention in solving the binding problem. *Neuron, 24*(1), 19–29, https://doi.org/10.1016/S0896-6273(00)80819-3.

Rosenholtz, R., Yu, D., & Keshvari, S. (2019). Challenges to pooling models of crowding: Implications for visual mechanisms. *Journal of Vision, 19*(7):15, 1–25, https://doi.org/10.1167/19.7.15.

Saarela, T. P., & Herzog, M. H. (2008). Time-course and surround modulation of contrast masking in human vision. *Journal of Vision, 8*(3), 23–23.

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*(2):7, 1–23, https://doi.org/10.1167/7.2.7.

Shechter, A., & Yashar, A. (2021). Mixture model investigation of the inner–outer asymmetry in visual crowding reveals a heavier weight towards the visual periphery. *Scientific Reports, 11*(1), 2116, https://doi.org/10.1038/s41598-021-81533-9.

Song, S., Levi, D. M., & Pelli, D. G. (2014). A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *Journal of Vision, 14*(5):3, 1–37, https://doi.org/10.1167/14.5.3.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11):8, 1024–1037, https://doi.org/10.1167/5.11.8.

Strasburger, H. (2020). Seven myths on crowding and peripheral vision. *i-Perception, 11*(3), 2041669520913052, https://doi.org/10.1177/2041669520913052.

Strasburger, H., & Malania, M. (2013). Source confusion is a major cause of crowding. *Journal of Vision, 13*(1):24, 1–20, https://doi.org/10.1167/13.1.24.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*(5):13, 1–82, https://doi.org/10.1167/11.5.13.

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

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136, https://doi.org/10.1016/0010-0285(80)90005-5.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14*(1), 107–141, https://doi.org/10.1016/0010-0285(82)90006-8.

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*(20), 2357–2369, https://doi.org/10.1016/S0042-6989(02)00197-9.

Wallace, J. M., Chung, S. T. L., & Tjan, B. S. (2017). Object crowding in age-related macular degeneration. *Journal of Vision, 17*(1):33, 1–13, https://doi.org/10.1167/17.1.33.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160–168, https://doi.org/10.1016/j.tics.2011.02.005.

Yashar, A., Chen, J., & Carrasco, M. (2015). Rapid and long-lasting reduction of crowding through training. *Journal of Vision, 15*(10):15, 1–15, https://doi.org/10.1167/15.10.15.

Yashar, A., Wu, X., Chen, J., & Carrasco, M. (2019). Crowding and binding: Not all feature dimensions behave in the same way. *Psychological Science, 30*(10), 1533–1546, https://doi.org/10.1177/0956797619870779.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396*(6706), 72–75, https://doi.org/10.1038/23936.

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *Journal of Vision, 10*(10):16, 1–12, https://doi.org/10.1167/10.10.16.

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

Zhu, Z., Fan, Z., & Fang, F. (2016). Two-stage perceptual learning to break visual crowding. *Journal of Vision, 16*(6):16, 1–12, https://doi.org/10.1167/16.6.16.