

VU Research Portal

Evolving the keys to visual crowding

Van der Burg, Erik; Olivers, Christian N.L.; Cass, John

published in

Journal of Experimental Psychology: Human Perception and Performance
2017

DOI (link to publisher)

[10.1037/xhp0000337](https://doi.org/10.1037/xhp0000337)

document version

Publisher's PDF, also known as Version of record

document license

Article 25fa Dutch Copyright Act

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Van der Burg, E., Olivers, C. N. L., & Cass, J. (2017). Evolving the keys to visual crowding. *Journal of Experimental Psychology: Human Perception and Performance*, 43(4), 690-699.
<https://doi.org/10.1037/xhp0000337>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Evolving the Keys to Visual Crowding

Erik Van der Burg
University of Sydney and Vrije Universiteit Amsterdam

Christian N. L. Olivers
Vrije Universiteit Amsterdam

John Cass
Western Sydney University

Peripheral vision can be severely impaired by nearby clutter. Decades of research using sparse displays have established that this phenomenon, known as visual crowding, follows Bouma's rule: Interference occurs for target-distractor separations up to half the target's eccentricity. Although considered a fundamental constraint on human vision, it is unclear whether Bouma's rule holds in dense heterogeneous visual environments. Using a genetic algorithm we investigated crowding in densely cluttered displays. Participants were instructed to identify the orientation of a target line (6° eccentricity) among 284 distractor lines. Displays supporting highest accuracy were selected ("survival of the fittest") and combined to create new displays. Performance improved over generations, predominantly driven by the emergence of horizontal flankers within 1° of the near-vertical target, but with no evidence of interference beyond this radius. We conclude that Bouma's rule does not necessarily hold in densely cluttered displays. Instead, a nearest-neighbor segmentation rule provides a better account.

Public Significance Statement

The greatest constraint on human vision is crowding: the inability to identify peripheral visual objects in cluttered environments. Crowding has been extensively investigated using sparse displays. Little is known, however, on the characteristics of crowding in complex, densely cluttered displays. Here we investigate crowding in cluttered visual scenes using a genetic algorithm, which "selects" large, initially random distractor arrays based on behavioral performance. Our final "evolved" stimuli show that the previously established rules believed to govern crowding are incomplete and that a nearest-neighbor rule provides a better account of human visual performance. We anticipate our approach could be adapted to a wide variety of tasks in cognitive neuroscience allowing investigation of large realistic stimulus search spaces.

Keywords: visual crowding, Bouma's rule, evolution

Supplemental materials: <http://dx.doi.org/10.1037/xhp0000337.supp>

Our ability to identify objects in our visual periphery may be severely corrupted by the mere presence of nearby flanking objects. This deleterious effect on visual performance, known as *visual crowding*, is most profound when flanking stimuli are presented within close proximity of the target, and reduces gradually with increasing separation until a critical limit of about half the target eccentricity is reached, beyond which performance is unimpaired (Bouma, 1970; see e.g., Strasburger, Harvey, & Rent-

schler, 1991; Whitney & Levi, 2011, for a review). This critical separation, known as *Bouma's rule*, is considered to be the defining characteristic of crowding (Pelli & Tillman, 2008), although the exact parameters may vary depending on the stimuli used (see Table 4 in Pelli, Palomares, & Majaj, 2004).

Visual crowding has been studied extensively over the past three decades. The majority of studies have used sparse displays, typically involving a target object surrounded by one, two, or four distractor objects (Andriessen & Bouma, 1976; Bouma, 1970; Kooi, Toet, Tripathy, & Levi, 1994; Toet & Levi, 1992). Figure 1a represents such a typical display.

Most natural visual environments, however, possess rather dense and heterogeneous image structure, which in principle seems capable of producing strong crowding (Wallis & Bex, 2012). So pervasive is crowding under natural viewing conditions that it has been suggested by several researchers to be *the* fundamental constraint on object recognition (Levi, 2008; Pelli & Tillman, 2008). One way the visual system copes with the deleterious effects of crowding is to use segmentation cues. If the target can be differ-

This article was published Online First February 9, 2017.

Erik Van der Burg, School of Psychology, University of Sydney, and Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam; Christian N. L. Olivers, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam; John Cass, School of Social Sciences and Psychology, Western Sydney University.

Correspondence concerning this article should be addressed to Erik Van der Burg, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Van der Boeorchstraat 1, 1181 BT Amsterdam, the Netherlands. E-mail: vanderburg.erik@gmail.com

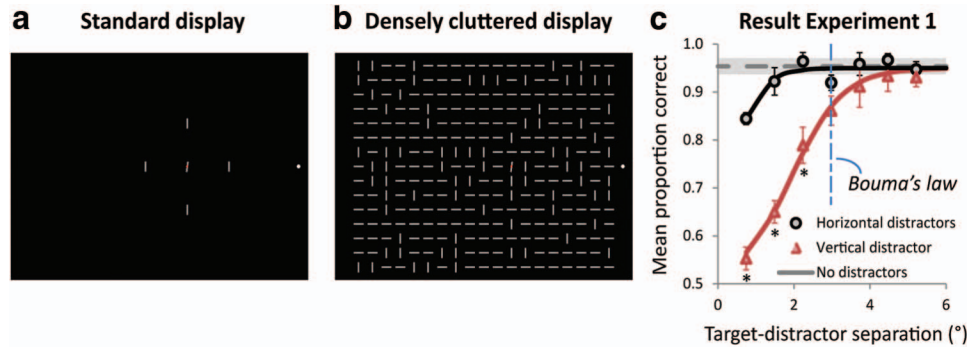


Figure 1. (a) Typical display to investigate visual crowding (Andriessen & Bouma, 1976). This set-up was used in Experiment 1. The target-distractor spacing as well as the orientation of the four distractors (all vertical or all horizontal) was manipulated. (b) Densely cluttered display used to investigate visual crowding in Experiment 2. Here the orientation of each of the 284 distractors (horizontal or vertical) was manipulated based on the participants' performance in combination with a genetic algorithm. On each trial target-distractor arrays could be presented either to the left or to the right of fixation (left visual field shown here). In both experiments participants were instructed to keep fixation at the center and report the orientation of the target line (tilted slightly clockwise or anticlockwise from vertical). Both panels are to scale. (c) Results of Experiment 1: Mean proportion of correct target identification as a function of distractor orientation and target-distractor separation. The red triangles signify trials in which the distractors were vertically oriented, whereas the black circles signify trials in which the distractors were horizontally oriented. The red and black continuous lines represent best fits (used for the Bouma simulation of the Experiment 2 data). The dotted line represents the condition in which no distractors were presented. Error bars represent the standard error of the mean (SEM). See the online article for the color version of this figure.

entiated from the flankers along some unique dimension (e.g., color), crowding effects become significantly weaker, and may even be obliterated altogether (Kooi et al., 1994). This “similarity principle” holds for numerous low-level features. By far the most studied of these is orientation. Indeed, whereas orientation acuity becomes significantly compromised by the mere presence of nearby flanking stimuli with similar orientation properties as the target (i.e., crowding), flankers that are orthogonally oriented with respect to the target exert little deleterious influence (Greenwood, Bex, & Dakin, 2010; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001).

Despite the ubiquity of dense structure in natural images (Bex, Solomon, & Dakin, 2009; Wallis & Bex, 2012), surprisingly few studies have ever investigated crowding in densely cluttered environments. Predictably, studies which have, have revealed severely degraded performance (Li, Thier, & Wehrhahn, 2000; Wallis & Bex, 2012; Williamson, Scolari, Jeong, Kim, & Awh, 2009). Interestingly, some notable exceptions to this general finding exist (see, e.g., Banks, Larson, & Prinzmetal, 1979; Manassi, Sayim, & Herzog, 2012; Pöder, 2006). Herzog, Manassi, and colleagues, for example, have published several articles showing that crowding caused by densely cluttered, but otherwise homogeneously arranged arrays of horizontal and vertical flanker lines, can be strongly attenuated by adding additional lines which promote high-level grouping between the flankers (Manassi et al., 2012; Manassi, Sayim, & Herzog, 2013). From a classical hierarchical view of the visual system this is theoretically significant as it implies both that crowding is not necessarily a purely feed-forward phenomenon, but one which is informed by higher-level grouping mechanisms (see Herzog & Manassi, 2015; Herzog, Sayim, Chich-erov, & Manassi, 2015, for a review).

In other studies, Dakin and colleagues (Dakin & Baruch, 2009; Robol, Casco, & Dakin, 2012) measured the effect of dense heterogeneous clutter on contour detection performance. Interestingly, they found that contour detection improved when flanking distractor elements most proximal to a near-collinear target path were perpendicularly oriented with respect to the Gabor elements along the path, as compared with when they were randomly oriented. This suggests that the brain may overcome the otherwise deleterious effects of crowding in densely cluttered scenes via mechanisms sensitive to local differences in orientation. Indeed, these authors show that a bank of second order filters sensitive to spatially adjacent orientation contrast are capable of yielding responses consistent with the aforementioned improvements in psychophysical performance.

In this study we use a novel method for investigating the brain's processing strategies for dealing with the problem of crowding in displays that are both highly cluttered and highly heterogeneous. We use a classic orientation discrimination task in which subjects identify whether a target line (6° eccentricity left or right of fixation) is tilted either 5° clockwise or anticlockwise from vertical. Each target was surrounded by a 19 × 15 rectangular array of 284 distractor lines (see Figure 1b for an example display), each assigned one of two possible orientation states (horizontal or vertical). Given the infeasibility of implementing a conventional factorial design to examine the effects of all possible 2²⁸⁴ stimulus configurations on target performance within large distractor arrays, we use an interactive genetic algorithm (GA; Van der Burg, Cass, Theeuwes, & Alais, 2015) to examine which distractors interfere with (or facilitate) target identification (see also Kong, Alais, & Van der Burg, 2016a, 2016b).

The application of GAs to the problem of crowding is unique. GAs involve optimization techniques that mimic natural selection (Holland, 1975). They start with a randomly sampled set of potential solutions to a complex problem. A fitness function evaluates the optimality of each solution in the set. Solutions that are poor are left to die out, while better solutions are allowed to generate offspring and propagate their advantageous traits over successive generations (i.e., “survival of the fittest”), using a cross-over technique. A small mutation rate is applied to avoid convergence to local minima, and to maintain diversity in a population (Holland, 1975). Although GAs are, in general, used to find an optimal solution for a complex problem, in our case we were not interested in finding an optimal display as such (as for vertical targets that is probably a display with only horizontal lines). Instead, we were interested in which specific distractors within the larger array affect performance most strongly (whether interfering or facilitating). Using this algorithm, we predict that distractors interfering with performance will, on average, evolve an orientation orthogonal to that of the vertical target (i.e., horizontal) in subsequent generations. Before that, however, we first established Bouma’s law for our experimental set-up, using sparse displays.

Experiment 1

Experiment 1 used sparse flanking displays and a standard full factorial design as conventionally used to investigate crowding, to demonstrate Bouma’s law in our experimental set-up. Either four horizontal or vertical distractors surrounded the target (see Figure 1a for an example), and the target-distractor separation was manipulated. A distractor absent condition was included as baseline. Based on the vast crowding literature, we expected to find increased visual crowding for vertical versus horizontal distractors, and that this pattern would follow Bouma’s law (Andriessen & Bouma, 1976; Parkes et al., 2001).

Method

Participants. Eleven participants (six female, mean age: 27.3 years, 22 to 42 years) participated in the experiment. Nine participants were VU University students and naïve as to the purpose of the experiment and paid €8 per hr. The other two participants were authors (EVdB and JC). Informed consent was obtained from each participant after the nature of the study was explained to them. The research was conducted with the ethical guidelines of the Faculty of Psychology and Education at the VU University, and those laid down in the Declaration of Helsinki.

Apparatus and stimuli. The experiment was run in a dimly lit cubicle and programmed by E-prime 2 software. Participants sat at a distance of approximately 70 cm from the LCD monitor (120 Hz refresh-rate). The display consisted of five white lines (0.43° length, 0.09° width). The four distractor lines were either all vertical or all horizontal, whereas the target line was tilted 5° either clockwise or anticlockwise from vertical. The distractors were presented above, below, on the left and right of the target, and the target-distractor separation was manipulated (0.749°, 1.49°, 2.24°, 2.98°, 3.73°, 4.47°, and 5.22°). A distractor absent condition was included as a baseline condition. The target and distractor lines were presented on either the left or right of a central white fixation dot (0.1° radius). The target eccentricity was always 5.96°. To

avoid source confusion between target and flanker locations, the two target locations were indicated by two red dots (0.02° radius). These dots were visible throughout the experiment. The background color was black and kept constant during the experiment.

Design and procedure. A trial started with the presentation of the fixation dot and the two red dots for a duration of 500 ms. Subsequently, the target-distractor array was either presented on the left or right of fixation for 150 ms, and then extinguished, leaving the fixation point and the two red dots signifying the potential target locations. Participants were asked to press the “z” or “m” key—unsped—to respectively indicate whether the target was tilted anticlockwise or clockwise of vertical. The next trial was initiated after the response. Target orientation, Target-distractor separation, and distractor orientation were balanced and randomly mixed within six experimental blocks of 96 trials each. Participants performed a single session.

Results

One participant was excluded from further analyses (see Experiment 2 for the exclusion criteria). The results of Experiment 1 are shown in Figure 1c. Here, the mean proportion of correct target identification is plotted as a function of distractor orientation, and target-distractor separation. An ANOVA on target identification performance was conducted with distractor orientation (vertical, or horizontal), and target-distractor separation as within subject variables. Here, and elsewhere in the manuscript, *p* values were Huynh-Feldt corrected to correct for sphericity violations. For each distractor orientation condition we fitted a four parameter sigmoid function (see Equation 1).

$$g(x) = y_0 + \frac{a}{1 + e^{\frac{x-x_0}{b}}} \quad (1)$$

Here, *x* represents the target-distractor separation. The intercept, *y*₀ was set to .5, as this corresponds to chance level. Parameter *a* was set to .45, as this corresponds to performance in the target alone condition minus chance level (i.e., upper asymptote). The slope, *b*, and parameter *x*₀ were free to vary. The black solid line in Figure 1c represents the optimal fit when the distractors are horizontally oriented (*r*² = .86; *b* = 0.67, and *x*₀ = 1.95), and the red solid line when the distractors are vertically oriented (*r*² = .99; *b* = 0.43, and *x*₀ = 0.24). Note that the parameters corresponding to the best fits were used for the Bouma simulation in Experiment 2.

Target identification accuracy. The ANOVA yielded a significant main effect of distractor orientation and target-distractor separation, *F*(1, 9) = 144.3, *p* < .00001, and *F*(6, 54) = 50.3, *p* < .00001, respectively. The two-way interaction was highly significant, *F*(6, 54) = 15.1, *p* < .00001, and further examined by two-tailed *t* tests. The *t* tests yielded a significant orientation effect when the target-distractor separation was 0.749, 1.49, 2.22°, *t*(9) = 6.7, *p* < .0001, *t*(9) = 8.2, *p* < .00005, and *t*(9) = 5.1, *p* = .001, respectively, and a trends toward a reliable distractor orientation effect when the target-distractor separation was 2.98, and 3.73°, *t*(9) *t*₀ = 1.9, *p* = .081, and *t*(9) = 2.0, *p* = .073, respectively. All other *t* tests yielded nonreliable effects: all *ps* > .147. Consistent with the literature, Experiment 1 delivered strong evidence for Bouma’s law (i.e., distractor interference for horizontal distractors

compared with vertical distractors when the target-distractor separation was smaller than half the target eccentricity) when using a conventional sparse crowding stimulus. Note that, here and elsewhere in the study we also conducted the same analyses on the arcsine transformed data, as proportional data may violate the equivariance assumptions of ANOVAs. These transformations and analyses yielded the same pattern of results and do not change the conclusions.

Experiment 2

The aim of Experiment 2 was to examine visual crowding in densely cluttered displays, using an evolutionary process shown in Figure 2a. For the first generation, 20 displays were generated by randomly assigning a horizontal or vertical orientation to each of the 284 distractors in the display with the constraint that the proportion of vertical distractors in the first generation was fixed such that overall performance for each participant approached 67% correct. Participants performed the orientation discrimination task for each of the randomly generated displays. The displays were repeated 12 times in random sequence, and a fit value (mean target identification accuracy) was assigned to each display. Following the survival of the fittest principle, the four “best” displays were then selected (i.e., those producing the highest fit value). These displays spawned a new generation of 12 displays by applying a cross-over and mutation technique. This cycle was repeated for a total of six generations.

We expected visual crowding to decrease over generations, as we apply a survival of the fittest strategy. Given that target-flanker similarity plays a crucial role in crowding (Andriessen & Bouma, 1976; Kooi et al., 1994) and that we employed a near vertical target, we predicted that with each generation some vertical distractors would likely evolve to become horizontal. Specifically, if visual crowding in dense displays is subject to Bouma’s law—as has been derived from using sparse displays (see Experiment 1)—then distractors within a radius of 3° (i.e., \sim half the target eccentricity) of the target should evolve to horizontal. If not, then

this questions the fundamental applicability of Bouma’s law in densely cluttered displays.

In addition to the experiment, we conducted two computer simulations using the same genetic algorithm as applied in the experiment. When analyzing the evolution of displays, one must be cautious in simply comparing each generation with the initial generation, as the proportion of verticals in the first generation may bias each subsequent generation. For instance, if the initial proportion of verticals is less than 50%, probabilistically the mutation rate will bias each subsequent generation to generate more verticals, until a probability of 50% is reached (i.e., the probability of mutating a given distractor to a horizontal or vertical distractor). We therefore added a simulation in which each generation was generated by randomly selected parents to generate a comparison condition, with initial conditions (proportion of horizontal to vertical) identical to those used in the experiment (we refer to this simulation as the *random simulation*). A second simulation was also conducted to investigate whether our GA (with just six generations and four sessions) can produce behavior consistent with Bouma’s law. For this we used identical GA parameter settings to those employed in the experiment, but now we used a fit function, corresponding to the behavioral Bouma-like functions observed in Experiment 1 (see Figure 1c), as criterion for selecting which displays to reproduce. We refer to this simulation as the *Bouma simulation*. Note that the evolution in the simulations and in the experiment is the same in all aspects, except that the selection of the four displays to reproduce differs.

Method

Participants. The same participants from Experiment 1 participated in Experiment 2 (Experiment 2 was done prior to Experiment 1).

Apparatus and stimuli. The experiment was the same as Experiment 1, except that the target was surrounded by 284 distractors (see Figure 1b for an example). Furthermore, instead of

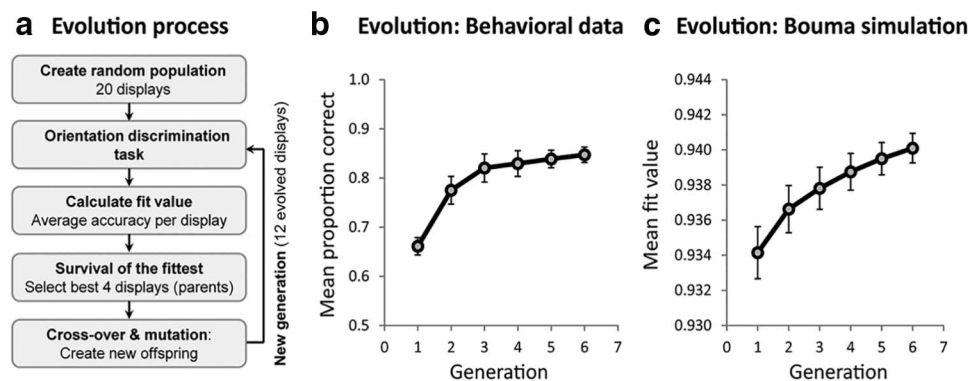


Figure 2. Evolution in Experiment 2. (a) The evolutionary process in Experiment 2. The genetic algorithm (GA) began with a first generation of 20 randomly determined displays. Those displays producing the highest mean accuracy were then selected to produce a new generation by applying a crossover and mutation technique, resulting in 12 evolved displays (the new generation). (b) Evolution in Experiment 2: Mean target identification accuracy as a function of generation. (c) Evolution in the Bouma simulation: Mean fit value (i.e., mean $g(x)$ per display) as a function of generation. Error bars represent the standard error of the mean (SEM).

using a factorial design, we used a GA to investigate visual crowding. The 284 distractor lines were either vertically or horizontally oriented, whereas the target line was tilted either 5° clockwise or anticlockwise from vertical. The line segments were presented on a 19 × 15 rectangular grid (13.4° horizontal and 10.4° vertical; 0.749° distance between elements) on either the left or right of the fixation dot. The target eccentricity was always 5.96°. To avoid source confusion between target and flanker locations, the two target locations were indicated by two red dots. These dots were visible throughout the experiment.

Design and procedure. Participants completed six blocks of trials, each block representing a generation in the evolution. Within the first block, 20 different display configurations (i.e., the population) were tested. For the other blocks (Generation 2–6), 12 different display configurations were tested. Each display was repeated 12 times (i.e., 240 trials for the first generation, 144 trials per block in Generation 2–6). On 50% of the trials, the array was presented on the left side of fixation, and on the remaining trials it was mirrored across the vertical axis of the screen so that it was presented on the right side of fixation. For the first generation, displays were generated by randomly assigning a horizontal or vertical orientation to each of the 284 distractors in the display with the constraint that the proportion of vertical distractors was fixed such that overall performance for each participant in the first generation approached 67% correct, in order to avoid floor and ceiling effects (which could compromise the evolution process). This proportion of verticals was estimated for each participant in a separate session prior to the experiment. In this session, participants were instructed to perform an orientation discrimination task identical to the one for the main experiment. The proportion of verticals was manipulated (.1, .3, .5, .7, or .9) and randomly mixed within two blocks of 40 trials each, and the point at which participants performed 67% correct was estimated from the function relating proportion correct to proportion verticals.

In the first generation participants performed the task 12 times on each of the 20 different displays. Each display (phenotype) was represented by its own gene (genotype), a 19 × 15 matrix whose gene position defined the location and orientation of each distractor in the display. A fit value was calculated for each display, which was defined as the mean proportion of correct target orientation identifications. Although the task was unspedded, mean reaction time (RT) on correct trials was used to determine the fitter one in case two displays yielded identical accuracy scores. Then, the best four displays (“parents”) were selected to reproduce (a “survival of fittest” principle). The four parents spawned a new generation of 12 displays (“children”) by applying a uniform cross-over procedure with a mixing ratio of 50%. As a result, the offspring were assigned approximately 50% of the genes from one parent and 50% of the genes from another parent. To establish uniform cross-over each bit in the parent genotype matrix was evaluated for exchange with a probability of 50%. Subsequently, each bit in the child genotype matrix (i.e., a distractor at a unique location) was assigned a 4% probability to randomly mutate to a vertical or horizontal distractor. These 12 new genotypes were each converted to their corresponding phenotype (i.e., a 284 item display), and as such constituted the next generation. This cycle was then repeated: Each display was presented 12 times and the four fittest displays were allowed to reproduce, through a total of six generations.

To make sure that participants maintained central fixation, we included two catch trials per block. In these catch trials, all distractors were vertically oriented, making it a virtually impossible task from central fixation, with an expected performance level to approach ~50% correct. In contrast, if participants were to fixate one of the target locations, then overall performance should approach 75% correct as they would reliably perceive the target’s orientation at that location, but be at chance level (50%) at the other target location. In total participants performed 3,840 experimental trials and 48 catch trials in four sessions. Participants received feedback about their overall proportion correct after each block.

Random simulation. In this simulation all GA parameter settings were identical to those used in the behavioral experiment (i.e., uniform cross-over; 4% mutation; six generations; 20 randomly created displays for Generation 1, and 12 displays for the remaining generations). In contrast to the GA experiment, the selection of the displays was randomly determined: Instead of on the basis of a fit-value, four random displays were selected for generating the next generation. Using the same software as for the experiment, we simulated 10 participants, each starting with exactly the same proportion of verticals as was set for each participant in the experiment. For example, if Participant 2 in the behavioral experiment started with .2 verticals in Generation 1, then the simulated Participant 2 started with the same proportion. For each simulated participant we simulated four sessions.

Bouma simulation. This simulation was the same as the random simulation in all aspects, except that we calculated the fit value for each display by using a *fit function*. This fit function was entirely based on the empirical data in Experiment 1 as we applied the parameters corresponding to the best fitting functions for each distractor orientation (see Figure 1c). For each distractor in a specific display we calculated $g(x)$, (Equation 1) based on the target-distractor separation, x , and the parameters corresponding to the distractor orientation. This was done for each distractor in the display and the *mean g(x)* was used as the fit value. Once, the fit values were calculated for all displays in the population, the four best displays (i.e., those displays with the highest *mean g(x)* score) were selected to reproduce.

Results

The data from one participant were excluded as this person showed considerable above chance performance on the catch trials (.64; group average = .53). The proportion of vertical distractors for the first generation across individuals ranged from .2 to .5, with a mean of .30. The overall proportion of vertical distractors obtained from the random simulation was subjected to an ANOVA with generation as the within subject variable. The ANOVA confirmed our suspicion that over generations the overall proportion of verticals increased (from on average .30 in the first generation to .34 in the final generation), $F(5, 54) = 24.4, p < .0001$, indicating that it is indeed important to compare the behavioral data and Bouma simulation with the random simulation baseline, as a significant increase of verticals across generations in the behavioral and Bouma simulation is less informative when partially driven by the mutation.

Mean proportion correct in the behavioral experiment. Figure 2b shows how overall identification performance improved

(i.e., crowding reduced) over generations. This was confirmed by an ANOVA on proportion correct with generation as the within subject variable, and which yielded a highly significant generation effect, $F(5, 45) = 30.7, p < .00001$.

Mean fit value in the Bouma simulation. Figure 2c reflects the evolution process in the Bouma simulation. An ANOVA on the mean fit value yielded a highly significant generation effect, $F(5, 45) = 68.8, p < .0001$, as the fit value increased over generations.

The important question is which regions of the display changed most, thus causing the observed decrease in crowding in the behavioral data. Specifically, we were interested in how this compared with whatever causes an increasing fit value in the Bouma simulation. The evolution of the distractors in the behavioral experiment and the Bouma simulation is shown in Movie 1 of the online supplemental materials (see Figure 3 for the final frame, corresponding to Generation 6). In Figure 3a, the proportion of vertical elements for each distractor in the

behavioral data (left panel) and the Bouma simulation (middle panel) is shown after subtraction of values of the random simulation. Note that in this figure, the target location is indicated by the white dot, and the fixation cross is presented on the right. Orange indicates the mean starting value. Lighter areas compared with orange signify an increase of vertical elements compared with the data derived by the random simulation, and darker areas signify a reduction. The difference between the behavioral and Bouma simulation is shown in Figure 3a (right panel). Here, orange indicates no difference between the behavioral data and the Bouma simulation. Lighter areas compared with orange signify more vertical elements in the behavioral data compared with the data derived by the Bouma simulation, and darker areas signify less vertical elements. Figure 3b is the same as Figure 3a, except that only significant differences are shown. The p values are shown in Figure 3c. Note that we did not correct for multiple comparisons in order

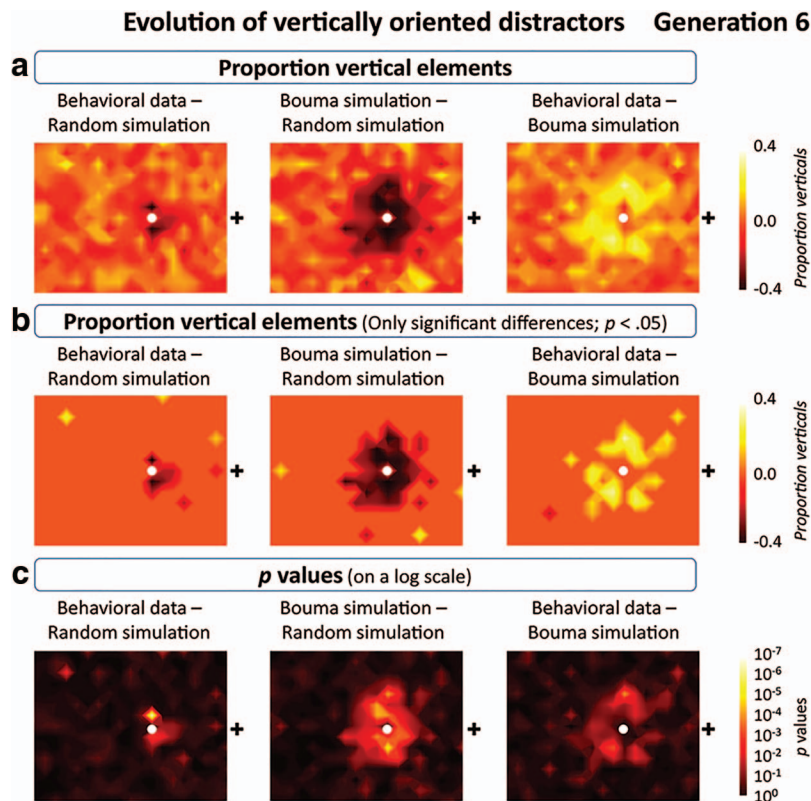


Figure 3. Evolution of the distractors in the behavioral experiment and the Bouma simulation for Generation 6. This figure corresponds to the final frame in Movie 1 of the supplementary materials. Note that the target location is indicated by the white dot, and the fixation cross is presented on the right. (a) Proportion of vertical elements for each distractor in Experiment 2 (behavioral data; left panel) and the Bouma simulation (middle panel) compared to the random simulation (all Generation 6). Orange indicates the mean starting value. Whereas lighter areas compared with orange signify an increase of vertical elements compared to the data derived by the random simulation, darker areas reflect a reduction. The difference in the proportion of verticals in the behavioral data minus Bouma simulation for Generation 6 is shown in the right panel. Here, orange indicates no difference between the behavioral data and the Bouma simulation. Lighter areas compared with orange signify more vertical elements in the behavioral data compared with the data derived by the Bouma simulation, and darker areas signify less vertical elements. (b) Same display as in Panel A, but only significant differences are plotted (two-tailed t tests with α set to .05; equal variances not assumed). The p values are shown in Panel C (on a log scale). See the online article for the color version of this figure.

to maximize the possibility of finding evidence for Bouma's law.

Spatial interference zones. In traditional sparse crowding displays, interference zones refer to the region around the target within flankers may exert their deleterious effects. In our displays, we assume interference zones correspond to locations that evolve (non-crowding) horizontal structure. As is evident from Movie 1, the distractors around the target become less vertical over generations in the behavioral data and the Bouma simulation (see darker areas in Figures 3a–c). However, the spatial extent of these interference zones seems to be much smaller for the behavioral data than for the data expected on the basis of Bouma's law (i.e., the Bouma simulation, see Figure 3c for the statistical difference). To estimate the spatial extent of these interference zones, for each distractor we calculated the target-distractor separation and binned the data in 19 bins of equal size (0.5° – 9.5° ; bin size: 0.5°). Figure 4a depicts the mean proportion of verticals for each bin (i.e., target-distractor separation) for each data set (behavioral, random simulation, and Bouma simulation) in Generation 6 (see Movie 2 in the online supplemental materials for the complete evolution of the spatial interference zones). We then conducted an ANOVA on the proportion of verticals with target-distractor separation as within subject variable and data set (behavioral, random simulation, and Bouma simulation) as between subject variable. The ANOVA yielded a reliable target-distractor separation effect, $F(18, 486) = 27.8, p < .00001$, and a significant data set effect, $F(1, 27) = 341.7, p < .00001$. The Data Set \times Target-Distractor Separation interaction was highly significant, $F(36, 486) = 12.9, p <$

.00001. Separate two-tailed t tests (equal variances not assumed) were conducted for each bin to investigate whether the three data-sets were significantly different from each other. The bars above Figure 4a illustrate the significant differences.

The data derived from the Bouma simulation supports the notion that one ought, in principle, to be able to find evidence for Bouma's law in cluttered displays using the same genetic algorithm as used in the behavioral experiment. Interestingly, similar spatial analyses of the evolved displays for the behavioral data set yielded no evidence for Bouma's law whatsoever. In this case, distractors within a narrow radius of $\sim 1^\circ$ of the target location interfered with performance, whereas others outside this range did not affect performance. In fact, the only difference between the data derived from the Bouma simulation, and the behavioral data (and also the random simulation) is the fit value used to evaluate, and select the displays for reproduction. In the case of Experiment 2, fit values corresponded to target identification performance, whereas for the Bouma simulation fit values were derived from the best fitting functions to the data in Experiment 1.

Anisotropies. Movie 1 (see behavioral data in Panels a–c; see also Figures 3a–c) suggests that crowding may be predominantly driven by the distractors *above* and *below* the target location, and to a lesser extent by the distractor toward fixation and away from fixation. To quantify this horizontal-vertical anisotropy we conducted an ANOVA on the proportion of vertical elements with position (distractors on the horizontal axis and distractors on the vertical axis) and target-distractor separation ($0.749^\circ, 1.49^\circ, 2.24^\circ, 2.98^\circ, 3.73^\circ,$

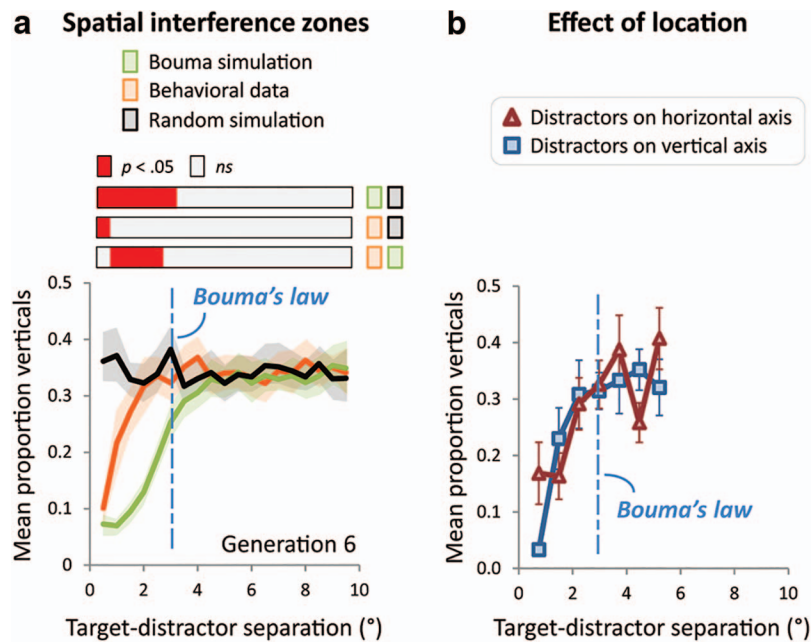


Figure 4. (a) Spatial interference zones for the behavioral data and the Bouma simulation compared with the random simulation for Generation 6 (final frame in Movie 2). Here the proportion of verticals is plotted for each bin (0.5° – 9.5° ; bin size: 0.5°) for the behavioral data (orange), the Bouma simulation (green), and the random simulation (black). See Supplementary Movie 2 for the full evolution process. The bars above the figure indicate whether the data sets were significantly different from each other (two-tailed t tests with α set to .05; equal variances not assumed). (b) Mean proportion of verticals as a function of target-distractor separation and position (vertical and horizontal axis) for Generation 6. Error bars represent the standard error of the mean (SEM). See the online article for the color version of this figure.

4.47°, and 5.22°). Figure 4b illustrates the proportion of verticals as a function of target-distractor separation and position for Generation 6. The ANOVA yielded a significant distance effect, $F(6, 54) = 9.7, p < .0001$, as the proportion of vertical elements increased with increasing distance (i.e., crowding). While the main effect of position was not reliable, $F(1, 9) = 2.2, p = .172$, the Distance \times Position interaction was significant, $F(6, 54) = 1.8, p = .025$. A close look at Figure 4b suggests that this interaction is driven by the distractors abutting the target (distance = 0.749°). Indeed, when we excluded this distance from the analyses, the Position \times Distance interaction was no longer significant, $F(5, 45) = 2.2, p = .086$. For the shortest distance (0.749°), a two-tailed t test was conducted to compare the proportion verticals along the vertical axis versus the horizontal axis. The t test yielded a significant effect, $t(9) = 2.5, p = .037$, as the proportion of verticals was significantly lower for the distractors above and below the target (.033) than on the left and right side of the target (.168).

Overall, our GA yielded a significant reduction in crowding over generations. Analyses of the spatial structure of the evolved displays indicates that this improvement in performance was caused by the emergence of horizontal distractors immediately above and below the near vertical target (i.e., at target end-zones), within $\sim 0.5^\circ$, and to a lesser extent, other distractors within a narrow radius of $\sim 1^\circ$ of the target location, and substantially less than the 3° expected on the basis of Bouma's law.

General Discussion

Using a genetic algorithm (Kong et al., 2016a, 2016b; Van der Burg et al., 2015) that selects and evolves displays based on target identification performance, we show that crowding weakens over successive generations. Spatial analyses of the evolved displays indicate that improvements in performance were in essence only caused by the distractors most proximal to the target, in particular by the emergence of orthogonally oriented distractors immediately above and below the target, and to a lesser extent, at positions within a narrow radius of $\sim 1^\circ$ of the target location. This result defies current theories of crowding which assume that the spatial extent of flanker interference ought to be no smaller than that prescribed by Bouma's law (Levi, 2008; Pelli & Tillman, 2008; Rosen, Chakravarthi, & Pelli, 2014), which in our set-up was 3° (as was also empirically confirmed in Experiment 1).

The emergence of flankers perpendicular to the target orientation, particularly at target end-zones is predicted by models which assume that crowding of line segments results from compulsory averaging of collinear flanking and target elements (Dakin, Cass, Greenwood, & Bex, 2010; Greenwood, Bex, & Dakin, 2009; Levi & Carney, 2009; May & Hess, 2008; Pöder, 2014). What is not specified in these "association field" models, however, is what influence, if any, flankers perpendicular to the target might have on visual crowding.

Most relevant to our study is work by Dakin and colleagues (Dakin & Baruch, 2009; Robol et al., 2012) who found that path detection performance in densely cluttered (and heterogeneously oriented environments) is strongly dependent upon the relative orientation of flanking elements adjacent to the target contour. Indeed, they found that: (a) perpendicular flanking elements adjacent to the path improved performance relative to randomly oriented flanking conditions, and conversely; (b) parallel flanking elements adjacent to the path impaired performance relative to

random conditions. Moreover, they showed that this pattern of performance is well-predicted by a set of filters sensitive to local orientation contrast. The perpendicular structure that evolved at target end-zones in our Experiment 2 is consistent with the predictions of this model.

Another possible explanation for our results (which may be consistent with the orientation contrast filter model of Dakin & Baruch, 2009; Robol et al., 2012) is that it is the distance of the target's nearest neighbor which determines whether crowding occurs or not in densely cluttered displays, rather than interelement distance per se. That is to say, when flanking elements most proximal to the target happen to be dissimilar to the target, performance is effectively released from any crowding that would otherwise be expected from more distant vertical elements falling within Bouma's limit. Indeed, our results show that when the target's nearest-neighbors are orthogonally oriented, performance in Generation 6 improved to "noncrowded" horizontally flanked baseline levels (85% compared with 84% correct in Experiments 1 and 2, respectively). Such a nearest-neighbor rule then also explains the crowding effects observed in the sparse arrays of the kind typically used to study crowding in which nearest-neighboring flankers are typically the only flanking elements in the display.

On the face of it, both the results of Experiment 2, and the nearest neighbor rule appear inconsistent with studies demonstrating improvements in identification performance (i.e., weakening of visual crowding) with increasing number of flankers, despite the proximity to the nearest-neighboring flankers remaining unchanged (Manassi et al., 2012, 2013; Pöder, 2006). Critically, however, in those studies reductions in crowding were due to the addition of extra flanking elements promoting perceptual grouping between the elements constituting the flanking array. If we assume that the critical nearest neighbor refers to the center of mass of the nearest "object" (either a single element or perceptually grouped set of elements), then crowding is expected to decrease, as adding more flankers increases the visual angle of the perceptual group, and thus the distance from the nearest neighbor to the target. Indeed, a recent study showed that the strength of crowding is not determined by the spatial proximity to a flanking shape's nearest constituent elements (e.g., its nearest edge), but to the central region of the shape (Rosen et al., 2014).

Relevant to such a nearest neighbor account is a study by Livne and Sagi (2010) who examined the effect of target-flanker separation in the presence (and absence) of flanking elements more proximal to the target. They found that the addition of an outer layer of flankers impaired performance relative to an inner-flanker alone condition. This result appears inconsistent with our nearest-neighbor account, which posits that the identity of the most proximate flanker determines whether crowding occurs or not. A possible explanation for the apparent discrepancy between our results and theirs' relates to differences in the aggregate disorder in each set of stimuli. The large, initially randomly (binary) oriented flanker arrays in our study ensured a high degree of disorder relative to the small, highly structured arrays used by Livne and Sagi (2010). Future research is necessary to determine what role, if any, stimulus heterogeneity (disorder) has on crowding.

A nearest neighbor account raises the intriguing question of how the emergence of flanking structure perpendicular and immediately adjacent to the target breaks crowding. One possibility is that

the strong orientation contrast evident in this stimulus arrangement makes the target appear perceptually more salient relative to the crowded conditions in which this orientation contrast signal is absent. This orientation salience account receives physiological support from a recent human fMRI study showing enhanced BOLD activation in extrastriate areas V3 and V4 for orientation “pop-out” stimuli (Bogler, Bode, & Haynes, 2013). More recent psychophysical evidence, however, indicates that target salience may not be a reliable predictor of crowding (Felisberti, Solomon, & Morgan, 2005).

Another possible explanation for the perpendicular target-flanker structures evolved in our displays is their resemblance to T-junctions. From a classical hierarchical view of the visual system (e.g., Wallis & Rolls, 1997) the evolution of T-junctions at the target location makes sense. Whereas primary visual cortex (V1) contains neurons, which encode both the orientation and location of individual lines and edges (Hubel & Wiesel, 1968), higher visual areas, such as V2 respond preferentially to line intersections, including T-junctions (Anzai, Peng, & Van Essen, 2007). Our results suggest, therefore, that the target has only to stimulate such a higher-level neuron to recover perceptual access to the orientation of the target object otherwise lost (crowded) due to lower-level interactions (Anderson, Dakin, Scharzwkopf, Rees, & Greenwood, 2012; Herzog & Manassi, 2015). This interpretation differs from the “high-level” grouping effects reported by Herzog, Manassi, and colleagues who showed that grouping between the flanking elements, but not the target, is sufficient to reduce crowding. By contrast our T-junction hypothesis predicts that in cluttered heterogenous displays grouping the target itself *with* its nearest (orthogonally oriented) flanking neighbors effectively immunizes the target from the deleterious effects of crowding. While the functional significance of our observation that T-junctions appear to break crowding in otherwise densely cluttered environments is unknown, we speculate that it might represent a processing strategy on the part of the visual system that promotes sensitivity to occlusion relationships in natural visual arrays.

In the present study we show that distractors within closest proximity of the target predominantly determine whether crowding occurs under cluttered circumstances. Distractors that are not immediate neighbors, but still fall within the area defined by Bouma’s law, play little role in crowding under these conditions. Our findings are better explained by a nearest neighbor rule, which states that crowding depends upon the identity of, and the distance to the target’s nearest neighboring elements. It is important to note that although the displays evolved using our *survival of the fittest* selection criterion imply that in densely cluttered scenes, the identity of the target’s nearest neighbors determine whether crowding occurs or not, this does not necessarily imply that the nearest neighbors are solely responsible for crowding when it does occur. That is to say, our effects are unidirectional: one cannot necessarily infer what displays are likely to produce the strongest crowding by simply reversing the branches of the algorithm. Indeed it is possible that more distant flankers may contribute to crowding in the absence of orthogonal nearest neighbors. Future studies could investigate this empirically using an analogous genetic algorithm to the one employed here but instead employing a “*survival of the sickest*” selection criterion.

References

- Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G., & Greenwood, J. A. (2012). The neural correlates of crowding-induced changes in appearance. *Current Biology*, *22*, 1199–1206. <http://dx.doi.org/10.1016/j.cub.2012.04.063>
- Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research*, *16*, 71–78. [http://dx.doi.org/10.1016/0042-6989\(76\)90078-X](http://dx.doi.org/10.1016/0042-6989(76)90078-X)
- Anzai, A., Peng, X., & Van Essen, D. C. (2007). Neurons in monkey visual area V2 encode combinations of orientations. *Nature Neuroscience*, *10*, 1313–1321. <http://dx.doi.org/10.1038/nn1975>
- Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics*, *25*, 447–456. <http://dx.doi.org/10.3758/BF03213822>
- Bex, P. J., Solomon, S. G., & Dakin, S. C. (2009). Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure. *Journal of Vision*, *9*, 1–19. <http://dx.doi.org/10.1167/9.1.1>
- Bogler, C., Bode, S., & Haynes, J.-D. (2013). Orientation pop-out processing in human visual cortex. *NeuroImage*, *81*, 73–80. <http://dx.doi.org/10.1016/j.neuroimage.2013.05.040>
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178. <http://dx.doi.org/10.1038/226177a0>
- Dakin, S. C., & Baruch, N. J. (2009). Context influences contour integration. *Journal of Vision*, *9*, 1–13. <http://dx.doi.org/10.1167/9.2.13>
- Dakin, S. C., Cass, J., Greenwood, J. A., & Bex, P. J. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision*, *10*, 14. <http://dx.doi.org/10.1167/10.10.14>
- Felisberti, F. M., Solomon, J. A., & Morgan, M. J. (2005). The role of target salience in crowding. *Perception*, *34*, 823–833. <http://dx.doi.org/10.1068/p5206>
- Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 13130–13135. <http://dx.doi.org/10.1073/pnas.0901352106>
- Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding changes appearance. *Current Biology*, *20*, 496–501. <http://dx.doi.org/10.1016/j.cub.2010.01.023>
- Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: A fresh look at object recognition. *Current Opinion in Behavioral Sciences*, *1*, 86–93. <http://dx.doi.org/10.1016/j.cobeha.2014.10.006>
- Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *Journal of Vision*, *15*, 5. <http://dx.doi.org/10.1167/15.6.5>
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Neurophysiology*, *195*, 215–243. <http://dx.doi.org/10.1113/jphysiol.1968.sp008455>
- Kong, G., Alais, D., & Van der Burg, E. (2016a). Competing distractors facilitate visual search in heterogeneous displays. *PLoS ONE*, *11*, e0160914. <http://dx.doi.org/10.1371/journal.pone.0160914>
- Kong, G., Alais, D., & Van der Burg, E. (2016b). An investigation of linear separability in visual search for colour suggests a role of recognisability. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1724–1738. <http://dx.doi.org/10.1037/xhp0000249>
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*, 255–279. <http://dx.doi.org/10.1163/156856894X00350>
- Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*, 635–654. <http://dx.doi.org/10.1016/j.visres.2007.12.009>

- Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology*, *19*, 1988–1993. <http://dx.doi.org/10.1016/j.cub.2009.09.056>
- Li, W., Thier, P., & Wehrhahn, C. (2000). Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *Journal of Neurophysiology*, *83*, 941–954.
- Livne, T., & Sagi, D. (2010). How do flankers' relations affect crowding? *Journal of Vision*, *10*, 1–14. <http://dx.doi.org/10.1167/10.3.1>
- Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision*, *12*, 13. <http://dx.doi.org/10.1167/12.10.13>
- Manassi, M., Sayim, B., & Herzog, M. H. (2013). When crowding of crowding leads to uncrowding. *Journal of Vision*, *13*, 10. <http://dx.doi.org/10.1167/13.13.10>
- May, K. A., & Hess, R. F. (2008). Effects of element separation and carrier wavelength on detection of snakes and ladders: Implications for models of contour integration. *Journal of Vision*, *8*, 1–23. <http://dx.doi.org/10.1167/8.13.4>
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*, 739–744. <http://dx.doi.org/10.1038/89532>
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*, 1136–1169. <http://dx.doi.org/10.1167/4.12.12>
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*, 1129–1135. <http://dx.doi.org/10.1038/nn.2187>
- Pöder, E. (2006). Crowding, feature integration, and two kinds of “attention”. *Journal of Vision*, *6*, 163–169. <http://dx.doi.org/10.1167/6.2.7>
- Pöder, E. (2014). Crowding by a single bar: Probing pattern recognition mechanisms in the visual periphery. *Journal of Vision*, *14*, 5. <http://dx.doi.org/10.1167/14.13.5>
- Robol, V., Casco, C., & Dakin, S. C. (2012). The role of crowding in contextual influences on contour integration. *Journal of Vision*, *12*, 3. <http://dx.doi.org/10.1167/12.7.3>
- Rosen, S., Chakravarthi, R., & Pelli, D. G. (2014). The Bouma law of crowding, revised: Critical spacing is equal across parts, not objects. *Journal of Vision*, *14*, 10. <http://dx.doi.org/10.1167/14.6.10>
- Strasburger, H., Harvey, L. O. Jr., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, *49*, 495–508. <http://dx.doi.org/10.3758/BF03212183>
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*, 1349–1357. [http://dx.doi.org/10.1016/0042-6989\(92\)90227-A](http://dx.doi.org/10.1016/0042-6989(92)90227-A)
- Van der Burg, E., Cass, J., Theeuwes, J., & Alais, D. (2015). Evolving the stimulus to fit the brain: A genetic algorithm reveals the brain's feature priorities in visual search. *Journal of Vision*, *15*, 8. <http://dx.doi.org/10.1167/15.2.8>
- Wallis, G., & Rolls, E. T. (1997). Invariant face and object recognition in the visual system. *Progress in Neurobiology*, *51*, 167–194. [http://dx.doi.org/10.1016/S0301-0082\(96\)00054-8](http://dx.doi.org/10.1016/S0301-0082(96)00054-8)
- Wallis, T. S. A., & Bex, P. J. (2012). Image correlates of crowding in natural scenes. *Journal of Vision*, *12*, 6. <http://dx.doi.org/10.1167/12.7.6>
- 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>
- Williamson, K., Scolari, M., Jeong, S., Kim, M. S., & Awh, E. (2009). Experience-dependent changes in the topography of visual crowding. *Journal of Vision*, *9*, 1–9. <http://dx.doi.org/10.1167/9.11.15>

Received August 5, 2015

Revision received October 10, 2016

Accepted October 12, 2016 ■

Instructions to Authors

For Instructions to Authors, please consult the January 2017 issue of the volume or visit www.apa.org/pubs/journals/xhp and click on the Manuscript Submission tab in the Journal Info box.