

Grouping, pooling, and when bigger is better in visual crowding

Mauro Manassi

Laboratory of Psychophysics, Brain Mind Institute,
Ecole Polytechnique Fédérale de Lausanne (EPFL),
1015 Lausanne, Switzerland



Bilge Sayim

Laboratory of Psychophysics, Brain Mind Institute,
Ecole Polytechnique Fédérale de Lausanne (EPFL),
1015 Lausanne, Switzerland;
Laboratoire Psychologie de la Perception (CNRS UMR
8158), Université Paris Descartes, Sorbonne Paris Cité,
France



Michael H. Herzog

Laboratory of Psychophysics, Brain Mind Institute,
Ecole Polytechnique Fédérale de Lausanne (EPFL),
1015 Lausanne, Switzerland



In crowding, perception of a target is strongly deteriorated by nearby elements. Crowding is often explained by pooling models predicting that adding flankers increases crowding. In contrast, the centroid hypothesis proposes that adding flankers decreases crowding—“bigger is better.” In foveal vision, we have recently shown that adding flankers can increase or decrease crowding depending on whether the target groups or ungroups from the flankers. We have further shown how configural effects, such as good and global Gestalt, determine crowding. Foveal and peripheral crowding do not always reveal the same characteristics. Here, we show that the very same grouping and Gestalt results of foveal vision are also found in the periphery. These results can neither be explained by simple pooling nor by centroid models. We discuss when bigger is better and how grouping might shape crowding.

Keywords: crowding, vernier acuity, perceptual organization, grouping

Citation: Manassi, M., Sayim, B., & Herzog, M. H. (2012) Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision*, 12(10):13, 1–14, <http://www.journalofvision.org/content/12/10/13>, doi:10.1167/12.10.13.

Introduction

In crowding, the perception of a target strongly deteriorates when flanked by neighboring elements. Crowding occurs both in foveal and peripheral vision. In crowding, it is often thought that the target per se is well visible but the discrimination of its features is impaired (Pelli, Palomares, & Majaj, 2004). For example, a letter is easily identified when presented alone. When flanking letters are added, the target letter is still visible, but its identification is impaired (Andriessen & Bouma, 1976; Levi, 2008; Pelli et al., 2004; Pelli & Tillman, 2008). Crowding occurs with all kinds of stimuli including high level stimuli, such as objects and faces (e.g., Louie, Bressler, & Whitney, 2007; Wallace & Tjan, 2011). Explanations of crowding can be found on various levels of description.

Neural mechanisms

Pooling models propose that stimuli are first analyzed by basic feature detectors, e.g., neurons in primary visual cortex V1, which then project to more complex feature detectors, e.g., neurons in higher visual areas with larger receptive fields (Freeman & Simoncelli, 2011; Levi, 2008; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli et al., 2004; Wilkinson, Wilson, & Ellemberg, 1997). Basic detectors signal the presence of the target (hence, target detection is not impaired). At the integration stage, feature location is lost since, for example, receptive fields are large and features are jumbled (Greenwood, Bex, & Dakin, 2009; Pelli et al., 2004) or inappropriate representations of image statistics are acquired due to saccade programming (Nandy & Tjan, 2012). In foveal vision, often lateral inhibition is proposed as a mechanism of interference by flankers (Westheimer & Hauske, 1975).

Stimulus layout

The window of crowding is often determined by Bouma's law, i.e., crowding occurs only when flankers are within a region of about 0.5 times the eccentricity of the target (Bouma, 1970; Pelli & Tillman, 2008). Bouma's law is usually expressed in terms of the distance between the outer contours of the target versus the inner contours of the flankers. Bouma's law is a description on the level of screen coordinates or stimulus layout. On the same level, Levi and Carney (2009) proposed that crowding strength is determined by the distance between the centroids of the flankers and the target, i.e., between the center of gravity of the target and the center of gravity of the flankers. Accordingly, increasing flanker size produces weaker crowding because flanker centroids move further away from the target.

Perceptual organization

Previous studies have shown that crowding decreases when flankers group (Banks, Larson, & Prinzmetal, 1979; Banks & White, 1984; Livne & Sagi, 2007; Wolford & Chambers, 1983). Malania, Herzog and Westheimer (2007) showed that foveal vernier offset discrimination strongly depends on the grouping of target and flankers. Single flankers that were either shorter, of the same length, or longer than the vernier deteriorated performance. Surprisingly, increasing the number of flankers decreased crowding in the shorter and the longer flanker condition but not in the same-length condition. Hence, adding flankers can, but does not have to, decrease crowding strength depending on the configuration of stimuli. Bigger can be better. Very similar results were found for Gabor stimuli in peripheral vision (Levi & Carney, 2009; Saarela, Sayim, Westheimer, & Herzog, 2009). These results cannot be explained by simple pooling models because more flankers increase noise and inhibition. Foveal grouping effects were further characterized. We showed that good Gestalt and the global layout of the stimulus determine crowding (Sayim, Westheimer, & Herzog, 2010, 2011). In these studies, we used vernier stimuli, which are perfectly suited for foveal crowding because of their small offsets and the small spacing between flankers and vernier.

It is debated whether foveal and peripheral crowding share the same mechanisms (Danilova & Bondarko, 2007; Levi, Hariharan, & Klein, 2002; Levi, Klein, & Hariharan, 2002; Pelli et al., 2004; Toet & Levi, 1992). In the first experiment, we show that adding more flankers can increase or decrease crowding, a finding similar to foveal results. In addition, we show that peripheral crowding with verniers exhibits the typical inward–outward asymmetry (Bouma, 1973). In the

second experiment, we show that figural manipulations affect peripheral crowding in the same manner as in foveal crowding. In the third experiment, we show that the global pattern of the entire stimulus configuration determines crowding.

Experiment 1A: When bigger is better

Adding flanking elements or increasing flanker size can reduce crowding both in foveal (Malania et al., 2007) and peripheral vision (Levi & Carney, 2009; Saarela et al., 2009). Simple pooling or lateral inhibition models predict the opposite result because more elements increase noise and inhibition. These results are in agreement with the centroid hypothesis that proposes that adding elements improves performance (Levi & Carney, 2009). However, adding flankers does usually not improve performance (e.g., Pelli et al., 2004; Wilkinson et al., 1997). Here, we show when more elements increase and decrease crowding.

Methods

We measured discrimination thresholds for vertical verniers presented alone or neighbored by various flanker configurations. Verniers consisted of two vertical lines slightly offset to the left or right. The task of the observers was to discriminate the vernier offset direction.

Apparatus and procedure

The experimental room was dimly illuminated (0.5 lx). Stimuli were generated on a Pentium-based computer and displayed on a high-resolution X-Y monitor (HP-1332A equipped with a P11 phosphor and controlled by a PC via fast 16-bit DA converters). Background luminance of the screen was below 1 cd/m². Luminance of stimuli was 80 cd/m². Viewing distance was 75 cm.

The vernier stimulus consisted of two vertical 40' (arcmin) long lines separated by a vertical gap of 4'. Flanker configurations were centered on the vernier stimulus and were symmetrical in the horizontal dimension. Each configuration was presented at an eccentricity of 3.88° to the right of a fixation cross (6' diameter). Eccentricity refers to the center of the target location. Observers were instructed to fixate the cross during the trial. In order to reduce uncertainty about the target position, we added two vertical lines (40' long), 100' above and below the center of the vernier.

Vernier and flankers were presented simultaneously for 150 ms.

Observers were asked to indicate the offset direction of the vernier by pressing one of two buttons. Auditory feedback was provided after incorrect or omitted responses. An adaptive staircase procedure (Taylor & Creelman, 1967) was used to determine the threshold for which observers reached 75% correct responses. Thresholds were determined after fitting a cumulative Gaussian to the data using probit and likelihood analyses. The starting offset was 16.66'. In order to avoid extremely large offsets, we restricted the adaptive procedure to 33.32' (i.e., twice the starting value).

After each trial, the screen remained blank for a maximum period of 3 s during which the observer was required to make a response. The screen was blank for 500 ms between the response and the next trial. In every block of 80 trials, the number of left and right offsets was balanced.

Stimulus configurations

An unflanked vernier was presented alone or was flanked by arrays of 1, 2, 4, or 8 vertical lines presented on each side (Figure 1). The directly neighboring lines were always placed at a horizontal distance of 23.33' from the vernier target. Interflanker spacing was also 23.33'. Three different flanker lengths were used: half (40'), equal (84'), or twice (168') the length of the vernier target. In a further condition, the short flankers were "jittered" along the vertical position (Figure 1d). For each number of flankers, the jittered configurations were the same in all trials for all observers. We will refer to these conditions as short, equal-length, long, and jittered flanker conditions. Each condition was presented in separate blocks of 80 trials. All conditions were measured twice (i.e., 160 trials) and randomized individually for each observer. To compensate for possible learning effects, the order of conditions was reversed after each condition had been measured once.

To relate threshold measurements to subjective ratings, we determined target-flankers grouping by asking 15 new observers (seven females) to judge how much the vernier stands out from the flankers on a scale from 0 (vernier does not stand out) to 10 (vernier strongly stands out). We plotted the results as 11 minus the standing-out ratings. High values indicate strong grouping between target and flankers (vernier does not stand out), low values indicate weak grouping between target and flankers (vernier stands out). Stimulus duration was unlimited. Otherwise flanker configurations were identical to the vernier offset discrimination experiments. As in the discrimination experiment, observers were asked to fixate the fixation cross during the trial.

Observers

Participants were paid students of the École Polytechnique Fédérale de Lausanne (EPFL). All had normal or corrected-to-normal vision, with a visual acuity of 1.0 (corresponding to 20/20) or better in at least one eye, measured with the Freiburg Visual Acuity Test (Bach, 1996). Observers were told that they could quit the experiment at any time they wished. Five observers (two females) performed the experiment with short, long, and jittered lines. Five other subjects (three females) participated in the equal-length flanker condition.

Statistics

We performed a regression analysis on the individual data, regressing threshold elevation against the log of the number of flankers. For each subject, this analysis yielded a slope and intercept of the regression line. We then performed *t*-tests to determine whether the slope of the regression lines differed significantly from 0.

Results

Results are plotted as threshold elevation relative to the unflanked vernier condition (Figure 1, upper panel). When the number of short flankers increased, performance improved (slope = 1.50, $t[4] = -4.60$, $p = 0.01$, Figure 1a). With equal-length flankers, performance remained constant (slope = -0.29, $t[4] = -0.37$, $p = 0.73$, Figure 1b). With the long flankers, we found a trend for facilitation (slope = -0.37, $t[4] = -2.15$, $p = 0.09$, Figure 1c). With jittered short flankers, performance strongly deteriorated (slope = 7.57, $t[4] = 3.75$, $p = 0.01$, Figure 1d).

Subjective ratings yielded similar results (Figure 1, lower panel). Target-flanker grouping is defined as 11 minus the standing-out ratings. Ratings decreased with small flankers (slope = -0.94, $t[14] = -2.28$, $p = 0.03$), were constant with equal-length flankers (slope = -0.16, $t[14] = -0.55$, $p = 0.59$), there was a trend for facilitation with long flankers (slope = -0.46, $t[14] = -1.98$, $p = 0.06$), and ratings increased with jittered short flankers (slope = 1.20, $t[14] = 4.62$, $p < 0.01$).

Discussion

Results of the equal-length and jittered flankers conditions are well in line with predictions of most pooling and inhibition models. More flankers increase noise and inhibition, and therefore, performance decreases or stays on a constant level. In the short line condition (Figure 1a), however, performance strongly improved when the number of lines increased although

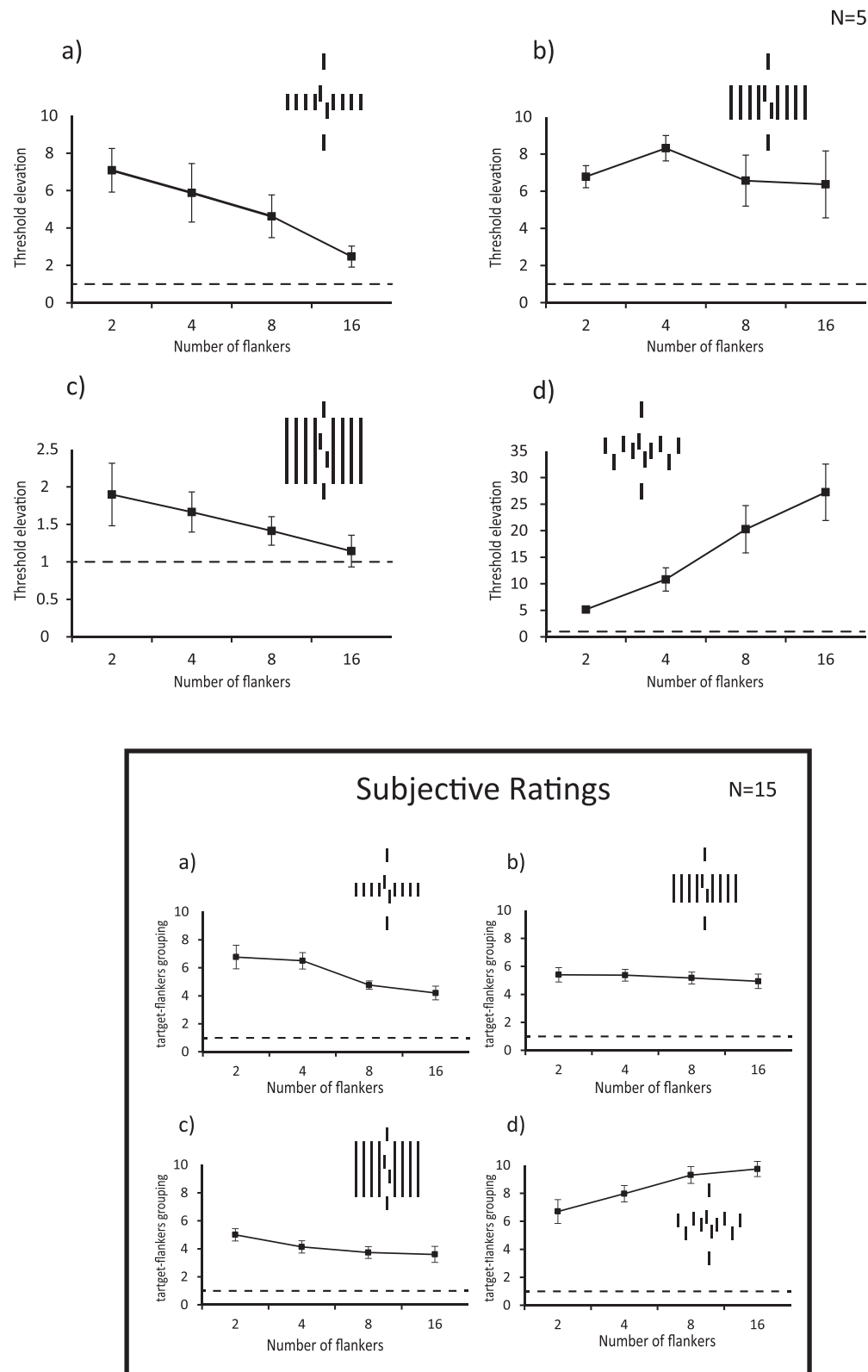


Figure 1. Upper panel: Results are plotted in terms of threshold elevation of the single vernier condition, i.e., thresholds of the flanking conditions divided by the threshold of the unflanked condition (dashed lines). Error bars indicate ± 1 standard error. A threshold elevation of 1.0 indicates no crowding; values larger than 1.0 indicate crowding. Performance improved when the number of short and long flankers increased (a and c), and stayed on a constant level for equal-length flankers (b). In contrast, performance deteriorated when the number of short, jittered flankers increased (d). In this condition, 16 flankers yielded a 26-fold increase in thresholds. Please note different scales of the y axis. Lower panel: Subjective ratings. Observers were asked to rate how much the vernier stands out from the flankers on a scale

→

it contained the same lines (at the same horizontal positions) as the jittered condition (Figure 1d). Simple pooling or other local interaction models cannot explain these results. With long flankers, performance improved as well but did not reach significance, possibly because performance for the single long flankers was already quite good.

We have shown that increasing the number or length of flankers can deteriorate performance, leave it unaffected, or improve it. Therefore, bigger is not always better, contrary to the centroid hypothesis, which proposes that performance improves when the number of flankers increases because the centroids of the flanker configurations move further away from the target. So, when is bigger better in crowding? We proposed that crowding diminishes when the target ungroups and stands out from the flankers (Malania et al., 2007; Saarela et al., 2009; Saarela, Westheimer, & Herzog, 2010; Sayim, Westheimer, & Herzog, 2008; Sayim et al., 2010, 2011). Bigger is better in the short and long flanker conditions because adding flankers enhances flanker–flanker grouping and, thus, target ungrouping from the flankers. Bigger is not better in the equal-length condition because grouping by similarity changes little by adding elements. Jittering the short lines gives rise to the percept of only one messy stimulus configuration in which the target vernier is embedded.

These considerations, even though speculative, correspond qualitatively well to the subjective ratings of vernier standing out (Figure 1, lower panel).

Experiment 1B: Inward–outward asymmetry

In peripheral vision, crowding is stronger when a flanker is presented away from the fovea than closer to it (Bouma, 1973). Here, we show that this inner–outer asymmetry occurs also with verniers.

The same apparatus and procedure were used as in Experiment 1A. In the first condition, only the unflanked vernier was presented. Next, the vernier was flanked by one vertical line either on the left or right side (i.e., inward and outward locations). Five observers (four females) participated in the experiment.

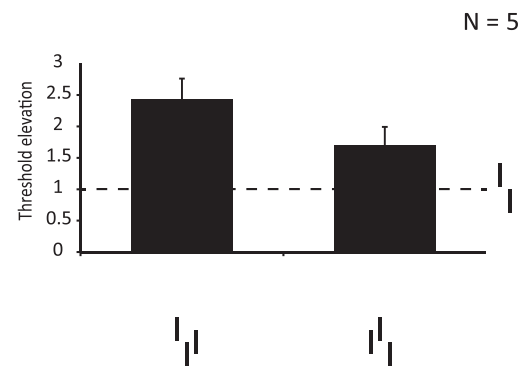


Figure 2. When adding one flanker on the right (outward location), performance deteriorated more strongly compared to when the flanker was added on the left (inward location). The dashed line indicates no threshold elevation (unflanked vernier).

When adding one flanker on the right (outward location), performance deteriorated more strongly compared to when the flanker was added on the left (inward location) ($t[4] = 3.23$, $p = 0.04$, Figure 2). These results are in agreement with the inward–outward anisotropy (Bouma, 1973; Petrov & Meleshkevich, 2011; Petrov, Popple, & McKee, 2007; Whitney & Levi, 2011).

Experiment 2: Good Gestalt

In the first experiment, we showed that adding flankers can decrease crowding when the target vernier ungroups from the flankers. In foveal vision, we showed that crowding also decreases when the vernier ungroups from the flankers by figural manipulations (Sayim et al., 2010). Here, we show that this holds true also in peripheral vision.

Methods

Apparatus and procedure were the same as in Experiment 1A. Observers were presented with five different stimulus configurations (Figure 3). An unflanked vernier served as a baseline condition. In the second condition, two vertical lines (84') were presented at a distance of 23.33' parallel to the vernier. In the third configuration, two rectangles were presented, with

from 0 (vernier does not stand out) to 10 (vernier strongly stands out). In order to compare psychophysical results with subjective ratings, results are plotted in reverse scale, i.e., standing-out ratings are subtracted from 11 and are called “target-flankers grouping,” with high values when target and flankers group, and low values when the vernier stands out from the flankers. Qualitatively, subjective ratings and discrimination thresholds are very similar. With short and long flankers, the vernier was perceived as standing out more and more from the configuration when the number of flankers increased. Standing out was roughly constant with same-length flankers. With jittered flankers, the vernier stood out less and less. The dashed line indicates standing out for the vernier alone condition (standard error is 0.0). Please note that y scales are different in discrimination but not in rating experiments.

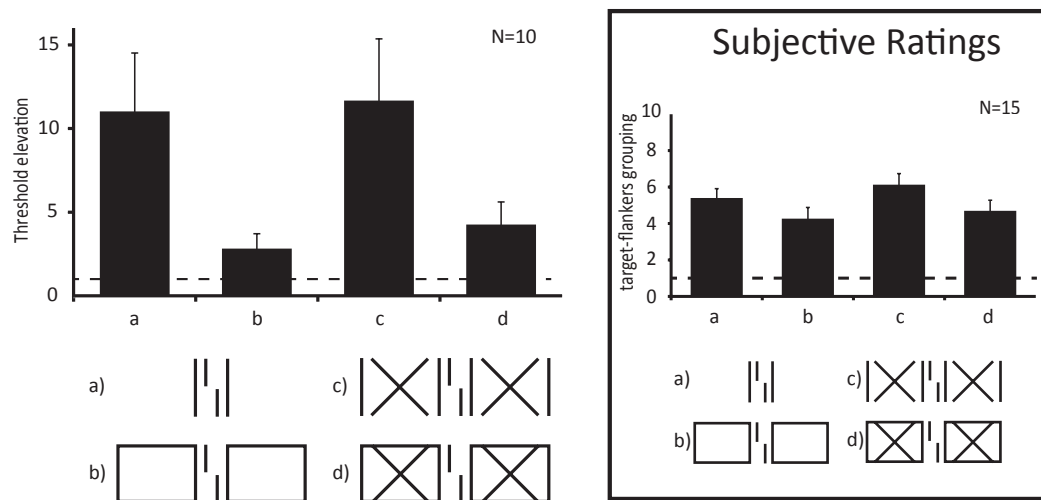


Figure 3. Left panel: Dashed line indicates no threshold elevation. Vernier offset discrimination deteriorated when adding two single lines (a). When the vernier was flanked by rectangles, there was an improvement in performance even though the flankers of (a) are part of the rectangles (b). Performance deteriorated when crossing the horizontal lines of the rectangles (the “energy” of the stimulus was the same, only the configuration had changed) (c). When adding crosses to the rectangles condition (b), performance improved (d). Flanker configurations (b, c, and d) have the same centroids. Right panel: Subjective ratings show a good qualitative match to discrimination thresholds.

one side corresponding to the flankers in the second condition and a width of $116'$. In the fourth condition, we shifted and rotated the two horizontal lines of the rectangle so that they formed a cross (angle of oblique lines is $45^\circ/135^\circ$, length is $116'$). The intersections coincided exactly with the centroid of the rectangles. The fifth configuration was identical to the previous condition, plus two horizontal lines were added.

Ten observers (five females) participated in the experiment. An iViewX-HiSpeed eye tracker from SensoMotoric Instruments (SMI) was used to control for eye movements for four subjects. The stimulus appeared only if subjects kept fixation. This group of subjects did not show different results compared to a group without eye tracking. Thus, data of the two groups were collapsed. As in Experiment 1A, seven new observers were asked to perform a rating task.

Statistics

Threshold data were analyzed with a repeated measures ANOVA. Tukey’s post-hoc tests were used for pairwise comparisons for all flanker configurations.

Results

We found a significant effect of flanker configuration on thresholds, $F(4, 36) = 7.64$, $p < 0.0001$ (Figure 3, left panel), and subjective ratings, $F(4, 56) = 25.38$, $p < 0.0001$ (Figure 3, right panel). When the vernier was flanked by two vertical lines, thresholds and ratings increased compared to the unflanked vernier condition

($p < 0.05$; Figure 3a, left and right panels). Thresholds decreased compared to the previous two-lines condition when the flanking lines formed a rectangle ($p < 0.05$, Figure 3b, left panel); ratings showed a similar pattern, but the difference was not significantly different (Figure 3b, right panel). When crossing the former upper and lower horizontal lines of the rectangles, thresholds and ratings increased compared to the previous condition ($p < 0.05$, Figure 3c, left and right panels). Finally, the configuration of rectangles plus crosses yielded lower thresholds and ratings, although with ratings the difference was not significantly different ($p < 0.05$, Figure 3d, left panel; $p > 0.05$, Figure 3d, right panel).

Discussion

The centroids were the same in the conditions of Figure 3b through d while performance strongly differed. However, according to the centroid hypothesis, performance should be identical in all three conditions (Figure 3b through d). Importantly, the single flankers in condition (a) were contained in all four stimulus configurations, ruling out models on local interactions. Hence, as in Experiment 1A, adding elements to flanker configurations can improve or deteriorate performance arguing against pooling and inhibition models.

We propose that the amount of grouping between target and flankers predicts best crowding strength. The single flankers of configuration (a) do not group with the vernier when they become part of a rectangle even though they are still present at the same location in

both conditions. Hence, models that take only local feature detector interactions into account fail to explain these results. For example, vertical line detectors sensitive to the flankers are equally activated in both conditions because the detectors are blind to the horizontal lines of the rectangles. Hence, crowding should be the same. However, this is not the case. Crowding depends on figural aspects in which the crowding power of single elements can be strongly modulated by their grouping into wholes, e.g., lines into rectangles. Whereas in [Experiment 1A](#), we showed how crowding varied when single flanking elements are grouped into wholes by similarity (i.e., length), here, we show how crowding varies when single elements are grouped into good Gestalts. These findings are very similar to what we found recently with foveal stimuli (Sayim et al., 2010).

Subjective reports on the vernier standing out ([Figure 3](#), right panel) correspond qualitatively well with the results on vernier discrimination ([Figure 3](#), left panel), although the ratings in some conditions do not differ significantly.

Experiment 3: Colors and patterns

In a previous study in foveal vision, we found that crowding increases when a red vernier is flanked by two red flankers (Sayim et al., 2008). Crowding was weak for two green flankers. However, crowding was very strong when these green flankers were flanked themselves by an array of alternating red and green flankers. These results show that crowding strongly depends on the global stimulus configuration. In the following experiment, we show that these results hold true also in peripheral vision.

Methods

Apparatus and procedure

Apparatus and procedure were the same as [Experiments 1](#) and [2](#) with the following changes. Stimuli were presented on a PHILIPS 201B4 CRT monitor driven by a standard accelerated graphics card. Screen resolution was set to 1024×768 pixels at 100Hz refresh rate. The white point of the monitor was adjusted to D65. The color space was linearized by applying individual gamma correction to each color channel. A Minolta CA-210 display color analyzer was used. Vernier and flankers consisted of (physically) isoluminant red and green lines presented on a black background. The luminance of the vernier target and the flankers was set to 10 cd/m^2 .

Stimulus configuration

Vernier and flankers were presented simultaneously for 100 ms. Stimuli were presented at an eccentricity of 5° (fixation dot diameter $3'$). All conditions were presented without the two vertical lines above and below the center of the vernier as used in [Experiments 1](#) and [2](#). The vernier was always red. We presented seven different flanker configurations. First, we added two red same-length flankers to the vernier ([Figure 4a](#)). Second, we added 10 red flankers on each side, increasing the length of the lines from the innermost flanker ($46'$) to the outermost ($82'$) to provide an indirect pointer to the target (butterfly conditions; [Figure 4b](#)). In the next two conditions, we changed the color of the flankers from red to green ([Figure 4c](#) and [d](#)). In the alternated configuration, the flankers were alternately red and green ([Figure 4e](#)). In the final two conditions, the green flankers or the red flankers, respectively, were removed from the alternated configuration ([Figure 4f](#) and [g](#)).

Observers

The Ishihara pseudoisochromatic color plates were used to test for red-green color vision deficiencies; no color deficiencies were observed in any participant. Four subjects (one female) participated, and the other seven subjects performed the subjective rating experiment.

Statistics

Threshold data were analyzed with a repeated measures ANOVA. Tukey's post-hoc tests were used for pairwise comparisons for all flanker configurations.

Results

Results are shown in [Figure 4](#) (upper and lower panels). We found a significant effect of flanker configuration on thresholds, $F(7, 21) = 24.86$, $p < 0.0001$ ([Figure 4](#), upper panel), and subjective ratings, $F(7, 98) = 23.72$, $p < 0.0001$ ([Figure 4](#), lower panel). Thresholds and ratings increased when the vernier was flanked by 2 or 10 red flankers compared to the unflanked condition ($p < 0.05$, [Figure 4a](#) and [b](#), upper and lower panels). With 2 or 10 green flankers, thresholds and ratings decreased considerably compared to the two previous conditions ($p < 0.05$, [Figure 4c](#) and [d](#), upper and lower panels). In the alternated butterfly configuration, thresholds and ratings again strongly increased ($p < 0.05$, [Figure 4e](#), upper and lower panels). In the threshold measurement, thresholds strongly decreased when only the red flankers of the alternated configuration were present ($p < 0.05$, [Figure 4f](#), upper panel). However, as an important exception, ratings

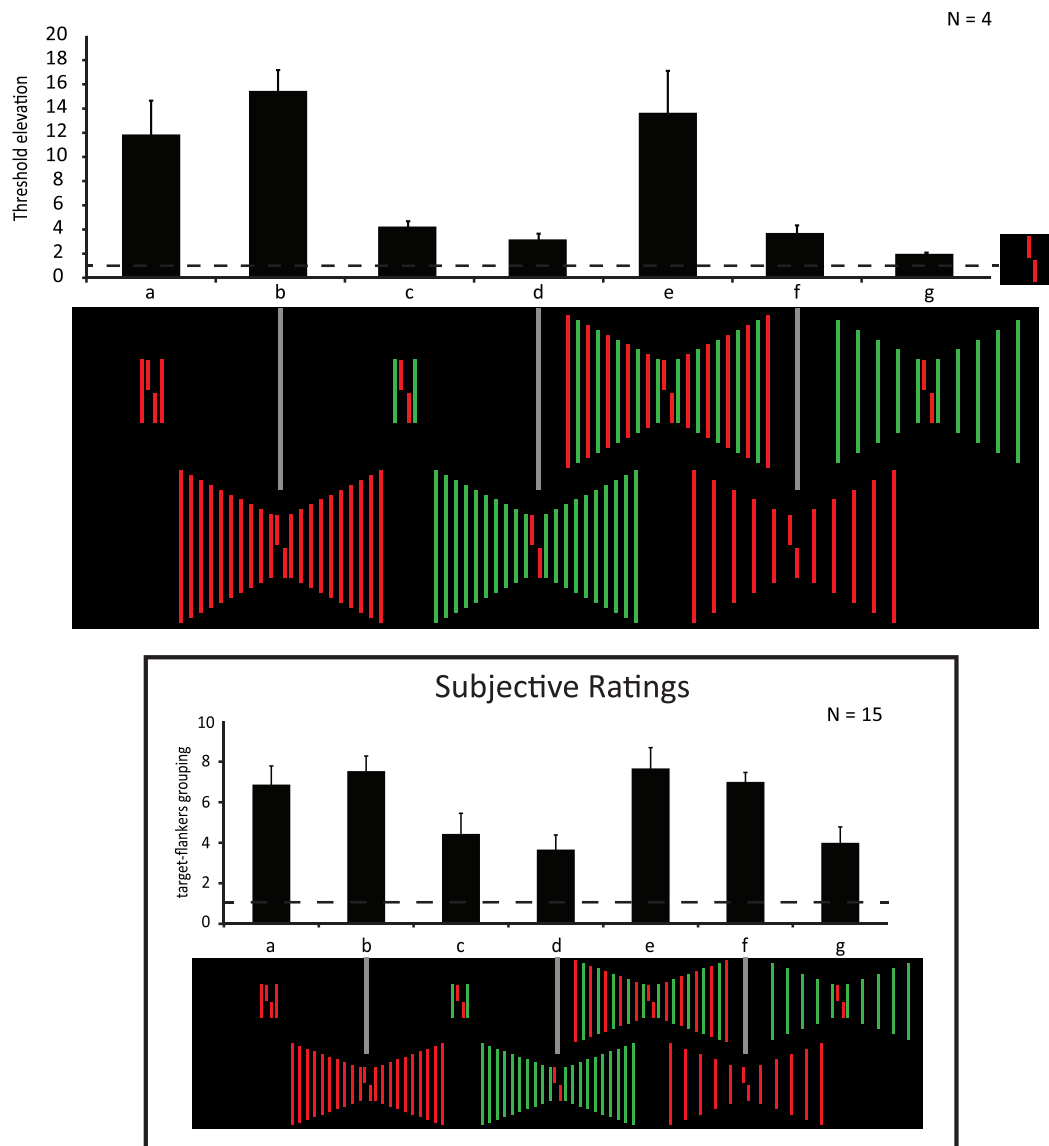


Figure 4. Upper panel: The dashed line indicates the threshold for the unflanked red vernier. Vernier offset discrimination deteriorated when adding 2 or 10 red flankers (a and b). Performance deteriorated much less when adding 2 or 10 green flankers (c and d). Performance deteriorated when adding a butterfly configuration with alternated colors (e). Performance improved when adding only green and red flankers (f and g). Lower panel: Subjective ratings on target-flankers grouping match well with discrimination thresholds, except for condition (f).

increased when the same condition was presented (Figure 4f, lower panel). Finally, thresholds and ratings decreased when only the green flankers were presented ($p < 0.05$; Figure 4g, upper and lower panels).

Discussion

Performance improved when changing the flanker color from red to green (Figure 4a through d). This result is well in line with previous findings showing that target-flanker similarity matters in crowding (Estes, 1982; Kooi, Toet, Tripathy, & Levi, 1994; Nazir, 1992).

The alternated configuration (Figure 4e) is the most interesting configuration because first, performance deterioration cannot be explained by local interactions of the vernier with the directly neighboring flankers. In the conditions of Figure 4d, e, and g, the vernier target was always flanked by the same two green flankers, but performance varied strongly. Second, the red flankers in the alternated condition cannot explain the deterioration because in the condition in Figure 4f performance was only slightly deteriorated. Hence, an additive effect of the monocular gratings cannot explain the deterioration in performance. It is the combination of red and green flankers.

We propose that regularity of the alternating lines yields grouping of the vernier with the flankers. Reading speed of alternating black–white letters is as slow as with either black or white letters, possibly, because letters group in all the three conditions (Chung & Mansfield, 2009). The centroid hypothesis cannot explain these results because in Figure 4b, d, e, and f, centroids of the flankers are at the same position but threshold elevations are substantially different.

Subjective ratings (Figure 4, lower panel) are qualitatively very similar to discrimination thresholds with the exception of Figure 4f. It seems that observers perceive the target vernier to group with the flankers. Still, thresholds are low because the spacing between flankers, particularly between target vernier and direct flankers, is wide. Alternatively, small receptive fields could have been used to discriminate the vernier offset (Figure 4, upper panel) and large receptive fields to perceive the global stimulus configuration (Figure 4, lower panel), thus yielding discrepancies between subjective ratings and threshold values.

General discussion

We have shown that grouping is a key feature of crowding. Crowding with verniers can yield remarkable threshold elevations larger than a factor of 26 (Figure 1d). Crowding with verniers shows the typical inward–outward asymmetry (however, such asymmetries are also found in masking, see Jiang & Chun, 2001). Our findings with peripheral vernier stimuli are very similar to previous findings with foveally presented verniers (Malania et al., 2007; Sayim et al., 2008, 2010). Very similar effects of grouping were also found with Gabor stimuli (Saarela et al., 2009; Saarela et al., 2010) and letters (Saarela et al., 2010) presented in the periphery.

Centroids

One of our main findings is that crowding can decrease, stay constant, or increase when the number of flankers or their length increases, i.e., the number and length of flankers per se does not predict crowding strength (Figure 1). When flankers were longer or shorter than the target vernier, crowding decreased when more flankers were presented (Figure 1a and c). When flankers were of the same length, crowding virtually did not change (Figure 1b). Crowding strongly increased when flankers were jittered (Figure 1d), even though crowding decreased for the very same flankers when not jittered (jittered vs. short condition). We previously obtained virtually the very same results with foveal stimuli (Malania et al., 2007). All other previous

studies have, to the best of our knowledge, shown that increases in flanker number increase crowding or leave it unchanged (Felisberti, Solomon, & Morgan, 2005; Parkes et al., 2001; Pelli et al., 2004; Pöder & Wagemans, 2007; Strasburger, Harvey, & Rentschler, 1991).

Based on the results in foveal vision, we proposed that grouping plays a key role in crowding (Malania et al., 2007). Crowding is weak or absent when the target ungroups from the flankers. Crowding increases when target and flankers group.

Levi and Carney (2009) also found that increases in flanker size decreases crowding. They proposed that bigger is better and explained their findings by the distance between the target and the centroids of the flankers (rather than between the inner contours of the flankers and the outer contours of the target, as in Bouma's law). However, increasing the size or number of flankers does not always lead to a reduction of crowding. The centroids in the short, equal, and long multi-flanker conditions of the present study are always the same, but performance varies nonmonotonically from short to equal to long flankers (Figure 1a through c). Likewise, the centroids are identical in the configurations of Figure 3b through d, but performance varies strongly. In addition, the centroids for the alternated color grating and the wider red and green gratings are nearly identical, but performance is clearly not (Figure 4d through f). Also in Levi and Carney (2009), performance did not improve when more than eight flankers were presented and their size increased. Hence, bigger is not always better. We propose that bigger is better when increases in the number or size of the flankers lead to ungrouping of the target from the flankers. Bigger is worse when it enhances grouping of the target with the flankers (Malania et al., 2007; Saarela et al., 2009; Saarela et al., 2010; Sayim et al., 2008, 2010, 2011).

According to the centroid hypothesis (Levi & Carney, 2009), the centroids of the flanker arrays to the right and left of the vernier determine crowding. It might be argued that this is only true when the vernier ungroups from the flankers. When vernier and flankers group, as for example, in the equal-length condition, there is only one grouped configuration, i.e., one center of gravity, and hence, performance deteriorates. However, such explanation presupposes a grouping operation, such as “flanker arrays to the left and right.” Hence, grouping plays a key role in explaining crowding strength.

Pooling

Bouma's law and the centroid hypothesis operate on the level of stimulus rendering or stimulus coordinates. The grouping account operates on the level of

perceptual organization. On the level of neural explanations, pooling is one of the key components. In standard pooling models of crowding (Pelli et al., 2004; Wilkinson et al., 1997), a visual stimulus is first processed at basic feature detectors, which then project to higher level feature detectors. Higher level detectors combine the outputs of a number of low level detectors and thus, position information of the features is lost. How exactly features are pooled is usually not exactly specified.

In very simple pooling models, features of nearby elements are pooled. Another option is that pooling occurs only amongst neurons that code for the same or similar features in accordance with the notion that crowding is weaker when target and flankers are dissimilar (Kooi et al., 1994; Saarela et al., 2009; Saarela et al., 2010; Sayim et al., 2008, 2010). Such pooling models can explain why short and long single flankers yield better performance than single, equal-length flankers (Parkes et al., 2001; van den Berg, Roerdink, & Cornelissen, 2010). However, such models fail to explain why performance improves when the number of long and short flankers increases. These multi-flanker array configurations contain the very same single flankers next to the vernier target. The same holds true in [Experiment 2](#), where single flankers become part of figures, i.e., good Gestalts. Low level feature detectors are triggered by these flankers in the same way whether or not they are part of good Gestalts. Hence, performance should not improve, but it does. Very similar results were found for Gabor stimuli in peripheral vision (Saarela et al., 2009) where multi-flanker configurations of long and short Gabor flankers led to better performance than equal-length flanker configurations.

One recent study showed that remote flankers, which crowd very little themselves, exert strong nonadditive increases of crowding when combined with flankers closer to the target (supercrowding; Vickery, Shim, Chakravarthi, Jiang, & Luedeman, 2009). In this respect, the results of [Experiments 1A](#) and [3](#) also show very similar, nonlinear crowding effects. In [Experiment 3](#), the wider spaced red and green arrays of flankers crowded very little, but their combination led to strong crowding by a factor of 14 ([Figure 4](#)).

Grouping

Only very few studies have investigated the role of grouping in crowding. In general, most research on crowding uses one target that is flanked by a pair of flankers. Multi-element or complex flankers were rarely used in the past even though the grouping idea is not new. Wolford and Chambers (1983) showed that when

flankers group, crowding is reduced (see also Banks & Prinzmetal, 1976).

Livne and Sagi (2007) proposed that *flanker–flanker* grouping determines crowding (see also Livne & Sagi, 2010). With the same set of stimuli, but in different configurations, Chakravarthi and Pelli (2011) showed evidence that only *local target–flanker* binding matters, i.e., the *global flanker–flanker* configuration does not matter—an interpretation countered by Livne and Sagi (2011). Our results with vernier stimuli favor an explanation in terms of global target–flanker grouping because adding flankers can weaken crowding, and good Gestalt and regularity matter ([Figures 1, 3, and 4](#); Malaria et al., 2007; Saarela et al., 2009; Sayim et al., 2010, 2011; but see Felisberti et al., 2005).

As we have shown, grouping plays a key role in crowding but is neither sufficient nor necessary. For example, observers grouped the red flankers in [Figure 4f](#); however, the flankers do not crowd because the spacing is too wide. In general, remote flankers can group with a target without crowding. Interestingly, remote flankers that group, but do not crowd, can still modulate target perception (see also Sayim & Cavanagh, 2011). Hence, grouping is not sufficient. Performance deteriorates when flanker luminance or contrast increases (Dombrowe, Hermens, Francis, & Herzog, 2009; Kooi et al., 1994), even though this leads to ungrouping by luminance/contrast dissimilarity. Hence, grouping is not necessary.

Gestalt rules explain grouping well for simple stimulus configurations but often fail when more than one Gestalt rule applies. Likewise with crowding. For example, proximity is a strong Gestalt cue and, indeed, performance usually improves when flankers are moved away from the target (in this respect, Bouma's law can be seen also as a description on the level of perceptual organization in addition to the level of stimulus layout). However, in peripheral vision, Saarela et al. (2010) showed that *decreasing* proximity between the target and the closest flankers *increased* crowding when the entire flanker configuration became regular by this manipulation. Hence, it seems that the Gestalt rule of good Gestalt or Pragnanz counteracted proximity. In addition, it is often unclear whether the Gestalt rules operate on a stimulus description level (e.g., proximity) or on a level of perceptual organization (good Gestalt). For this reason, we propose that subjective measurements of grouping must be determined in addition to accuracy measurements.

Grouping is a key factor in crowding. However, grouping itself cannot explain why performance deteriorates when target and flankers group. Grouping needs to be linked to neural mechanisms such as inhibition, feature pooling, or limited attentional resolution (He, Cavanagh, & Intriligator, 1996; Strasburger, 2005). Mechanisms are needed explaining why,

for example, (location) information of features is *lost* (Huckauf & Heller, 2002). Thus, how can grouping be linked to neural mechanisms? As mentioned, simple pooling or inhibition models cannot explain the results presented here.

It may well be that grouping precedes crowding, i.e., the visual system first groups elements into wholes and then within each group a crowding mechanism operates (Parkes et al., 2001). However, why should this be the case? Feature pooling across the visual field can be explained by receptive fields properties. Pooling within groups cannot because grouping between elements can easily change while receptive fields do not. Models are needed that take flexible element bindings into account.

May and Hess (2007) proposed such a model in which two, rather than one, mechanisms explain contour integration and crowding. A collinear and a noncollinear binding mechanism compete about the ownership of elements. Grouping of an element to one group can ungroup the elements from another group. This model also explained crowding of letters well. Whereas such a model can possibly not be directly applied to our vernier stimuli, the model provides a framework of how global aspects of crowding emerge by the competition of two mechanisms. Our results and the results of contour integration also show strong phenomenological similarities. Collinear Gabor “snakes” are perceived as clearly different from flanking Gabors. Likewise, the single flankers in Figure 3 are not seen as flankers but as a part of a rectangle—and they do not group with the vernier anymore. However, the exact characteristics of the underlying neural computations remain largely unknown for the moment.

Another question is *when* crowding occurs. Crowding may occur as a consequence of grouping or *during* grouping. For example, models with recurrent lateral inhibition have been proven to explain a large variety of grouping effects in backward masking that show the very same characteristics reported here in crowding (Hermens, Luksys, Gerstner, Herzog, & Ernst, 2008; Hermens, Scharnowski, & Herzog, 2009). In these models, representations of remote flankers, even when not connected directly, can influence each other via intermediate flanker representations in a time consuming manner (Hermens et al., 2008; Herzog, Ernst, Etzold, & Eurich, 2003). Applied to the stimuli of this contribution, activity of the vernier is inhibited for equal-length flankers because the vernier is in the center of a redundant structure. In long and short multi-flanker configurations, the flankers inhibit each other but not the vernier, and hence performance is better than in equal-length conditions.

Future work has to link descriptions on the level of perceptual organization (grouping) with explanations on the neural level (recurrent inhibition). For the

moment, we propose that one of the best predictors of the strength of crowding is the extent to which target and flanking elements are grouped. However, the ultimate goal is to provide a model that links all levels of explanations.

Acknowledgments

We would like to thank Keith May and two anonymous reviewers for their useful comments on this manuscript. We thank Marc Repnow for technical support. This work was supported by the Swiss National Science Foundation (SNF) Project “Basics of visual processing: What crowds in crowding?”

Commercial relationships: none.

Corresponding author: Mauro Manassi.

Email: mauro.manassi@epfl.ch.

Address: Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland.

References

- Andriessen, J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research*, 16(1), 71–78.
- Bach, M. (1996). The Freiburg Visual Acuity test—automatic measurement of visual acuity. *Optometry & Vision Science*, 73(1), 49–53. [PubMed]
- Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics*, 25(6), 447–456. [PubMed]
- Banks, W. P., & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception & Psychophysics*, 19(4), 361–367.
- Banks, W. P., & White, H. (1984). Lateral interference and perceptual grouping in visual detection. *Perception & Psychophysics*, 36(3), 285–295. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226(5241), 177–178. [PubMed]
- Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, 13(4), 767–782. [PubMed]
- Chakravarthi, R., & Pelli, D. G. (2011). The same binding in contour integration and crowding. *Journal of Vision*, 11(8):10, 1–12, <http://www.journalofvision.org/content/11/8/10>, doi:10.1167/11.8.10. [PubMed] [Article]

- Chung, S. T., & Mansfield, J. S. (2009). Contrast polarity differences reduce crowding but do not benefit reading performance in peripheral vision. *Vision Research*, 49(23), 2782–2789.
- Danilova, M. V., & Bondarko, V. M. (2007). Foveal contour interactions and crowding effects at the resolution limit of the visual system. *Journal of Vision*, 7(2):25, 1–18, <http://www.journalofvision.org/content/7/2/25>, doi:10.1167/7.2.25. [PubMed] [Article]
- Dombrowe, I., Hermens, F., Francis, G., & Herzog, M. H. (2009). The roles of mask luminance and perceptual grouping in visual backward masking. *Journal of Vision*, 9(11):22, 1–11, <http://journalofvision.org/9/11/22>, doi:10.1167/9.11.22. [PubMed] [Article]
- Estes, W. K. (1982). Similarity-related channel interactions in visual processing. *Journal of Experimental Psychology: Human Perception & Performance*, 8(3), 353–382. [PubMed]
- Felisberti, F. M., Solomon, J. A., & Morgan, M. J. (2005). The role of target salience in crowding. *Perception*, 34(7), 823–833. [PubMed]
- Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, 14(9), 1195–1201. [PubMed]
- 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 the United States of America*, 106(31), 13130–13135. [PubMed]
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, 383(6598), 334–337. [PubMed]
- Hermens, F., Luksys, G., Gerstner, W., Herzog, M. H., & Ernst, U. (2008). Modeling spatial and temporal aspects of visual backward masking. *Psychology Review*, 115(1), 83–100. [PubMed]
- Hermens, F., Scharnowski, F., & Herzog, M. H. (2009). Spatial grouping determines temporal integration. *Journal of Experimental Psychology: Human Perception & Performance*, 35(3), 595–610. [PubMed]
- Herzog, M. H., Ernst, U. A., Etzold, A., & Eurich, C. W. (2003). Local interactions in neural networks explain global effects in gestalt processing and masking. *Journal of Neural Computation*, 15(9), 2091–2113. [PubMed]
- Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. *Visual Cognition*, 9(2), 889–910.
- Jiang, Y., & Chun, M. M. (2001). Asymmetric object substitution masking. *Journal of Experimental Psychology: Human Perception & Performance*, 27(4), 895–918. [PubMed]
- 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(2), 255–279. [PubMed]
- Levi, D. M. (2008). Crowding—an essential bottleneck for object recognition: A mini-review. *Vision Research*, 48(5), 635–654. [PubMed]
- Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology*, 19 (23), 1988–1993. [PubMed]
- Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, 2(2):3, 167–177, <http://www.journalofvision.org/content/2/2/3>, doi:10.1167/2.2.3. [PubMed] [Article]
- Levi, D. M., Klein, S. A., & Hariharan, S. (2002). Suppressive and facilitatory spatial interactions in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision*, 2(2):2, 140–166, <http://www.journalofvision.org/content/2/2/2>, doi:10.1167/2.2.2. [PubMed] [Article]
- Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision*, 7(2):4, 1–12, <http://www.journalofvision.org/content/7/2/4>, doi:10.1167/7.2.4. [PubMed] [Article]
- Livne, T., & Sagi, D. (2010). How do flankers' relations affect crowding? *Journal of Vision*, 10(3):1, 1–14, <http://journalofvision.org/10/3/1>, doi:10.1167/10.3.1. [PubMed] [Article]
- Livne, T., & Sagi, D. (2011). Multiple levels of orientation anisotropy in crowding with gabor ankers. *Journal of Vision*, 11(13):18, 1–10, <http://www.journalofvision.org/content/11/13/18>, doi:10.1167/11.13.18. [PubMed] [Article]
- Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision*, 7(2):24, 1–11, <http://www.journalofvision.org/content/7/2/24>, doi:10.1167/7.2.24. [PubMed] [Article]
- Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *Journal of Vision*, 7(2):1, 1–7, <http://www.journalofvision.org/content/7/2/1>, doi:10.1167/7.2.1. [PubMed] [Article]
- May K. A., & Hess, R. F. (2007). Ladder contours are undetectable in the periphery: A crowding effect?

- Journal of Vision*, 7(13):9, 1–15, <http://www.journalofvision.org/content/7/13/9>, doi:10.1167/7.13.9. [PubMed] [Article]
- Nandy, A. S., & Tjan, B. S. (2012). Saccade-confounded image statistics explain visual crowding. *Nature Neuroscience*, 15(3), 463–469. [PubMed]
- Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Research*, 32(4), 771–777. [PubMed]
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4(7), 739–744. [PubMed]
- 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, <http://www.journalofvision.org/content/4/12/12>, doi:10.1167/4.12.12. [PubMed] [Article]
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11(10), 1129–1135. [PubMed]
- Petrov, Y., & Meleshkevich, O. (2011). Locus of spatial attention determines inward–outward anisotropy in crowding. *Journal of Vision*, 11(4):1, 1–11, <http://www.journalofvision.org/content/11/4/1>, doi:10.1167/11.4.1. [PubMed] [Article]
- 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, <http://www.journalofvision.org/content/7/2/12>, doi:10.1167/7.2.12. [PubMed] [Article]
- Pöder, E., & Wagemans, J. (2007). Crowding with conjunctions of simple features. *Journal of Vision*, 7(2):23, 1–12, <http://www.journalofvision.org/content/7/2/23>, doi:10.1167/7.2.23. [PubMed] [Article]
- Saarela, T. P., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, 9(2):5, 1–11, <http://www.journalofvision.org/content/9/2/5>, doi:10.1167/9.2.5. [PubMed] [Article]
- Saarela, T. P., Westheimer, G., & Herzog, M. H. (2010). The effect of spacing regularity on visual crowding. *Journal of Vision*, 10(10):17, 1–7, <http://www.journalofvision.org/content/10/10/17>, doi:10.1167/10.10.17. [PubMed] [Article]
- Sayim, B., & Cavanagh, P. (2011). Long range grouping affects crowding. *Perception*, 40, 35.
- Sayim, B., Westheimer, G., & Herzog, M. H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. *Journal of Vision*, 8(8):12, 1–9, <http://www.journalofvision.org/content/8/8/12>, doi:10.1167/8.8.12. [PubMed] [Article]
- Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science*, 21(5), 641–644. [PubMed]
- Sayim, B., Westheimer, G., & Herzog, M. H. (2011). Quantifying target conspicuity in contextual modulation by visual search. *Journal of Vision*, 11(1):6, 1–11, <http://www.journalofvision.org/content/11/1/6>, doi:10.1167/11.1.6. [PubMed] [Article]
- Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, 5(11):8, 1024–1037, <http://www.journalofvision.org/content/5/11/8>, doi:10.1167/5.11.8. [PubMed] [Article]
- Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, 49(6), 495–508. [PubMed]
- Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America*, 41(4), 782–787.
- Toet, A., & Levi, D. M. (1992, Jul). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32(7), 1349–1357. [PubMed]
- van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology*, 6(1), e1000646. [PubMed]
- Vickery, T. J., Shim, W. M., Chakravarthi, R., Jiang, Y. V., & Luedeman, R. (2009). Supercrowding: Weakly masking a target expands the range of crowding. *Journal of Vision*, 9(2):12, 1–15, <http://www.journalofvision.org/content/9/2/12>, doi:10.1167/9.2.12. [PubMed] [Article]
- Wallace, J. M., & Tjan, B. S. (2011). Object crowding. *Journal of Vision*, 11(6):19, 1–17, <http://www.journalofvision.org/content/11/6/19>, doi:10.1167/11.6.19. [PubMed] [Article]
- Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research*, 15, 1137–1141. [PubMed]
- Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Science*, 15(4), 160–168. [PubMed]
- Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture

arrays. *Journal of the Optical Society of America A: Optics, Image Science, & Vision*, 14(9), 2057–2068. [[PubMed](#)]

Wolford, G., & Chambers, L. (1983). Lateral masking as a function of spacing. *Perception & Psychophysics*, 33(2), 129–138. [[PubMed](#)]