When crowding of crowding leads to uncrowding

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

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

Michael H. Herzog
Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

In object recognition, features are thought to be processed in a hierarchical fashion from low-level analysis (edges and lines) to complex figural processing (shapes and objects). Here, we show that figural processing determines low-level processing. Vernier offset discrimination strongly deteriorated when we embedded a vernier in a square. This is a classic crowding effect. Surprisingly, crowding almost disappeared when additional squares were added. We propose that figural interactions between the squares precede low-level suppression of the vernier by the single square, contrary to hierarchical models of object recognition.

Introduction

In object recognition, features are thought to be processed in a hierarchical, feedforward fashion in which low-level feature analysis, e.g., of edges and lines, precedes the analysis of complex features, such as shapes and objects (e.g., DiCarlo, Zoccolan, & Rust, 2012; Hubel & Wiesel, 1959, 1962; Riesenhuber & Poggio, 1999; Thorpe, Fize, & Marlot, 1996). For example, the processing of a square presupposes the analysis of its constituting four lines (Figure 1).

In crowding, the perception of a target strongly deteriorates when flanked by neighboring elements. Crowding is often explained by pooling or substitution models, which are well in the spirit of the hierarchical, feedforward model of object recognition. In pooling models, neurons in higher visual areas with larger receptive fields pool information from lower-level neurons with smaller receptive fields (Greenwood, Bex, & Dakin, 2010; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli, Palomares, & Majaj, 2004; Wilkinson, Wilson, & Ellemberg, 1997). Because of pooling, features of the target and the flankers are integrated, and, thus, feature identification is lost. In substitution models, because of positional uncertainty (Huckauf & Heller, 2002; Strasburger, Harvey, & Rentschler, 1991) or limited attentional resolution (Strasburger, 2005), features of the target and flankers are mislocalized or not "accessible" by attention (He, Cavanagh, & Intriligator, 1996).

A prediction of all of these models is that, when adding more flankers, crowding increases because more noise is pooled or more elements can be confused. However, this is not always the case (Manassi, Sayim, & Herzog, 2012; Wolford & Chambers, 1983). For example, in a previous contribution, we determined vernier offset discrimination with different flanker configurations (Figure 2). In the first experiment, crowding was strong when the flankers had the same length as the vernier (Figure 2a). Increasing or decreasing the length of the flankers led to a *decrease* in crowding compared to the previous condition (Figure 2b and c). Pooling and substitution models can explain the improvement in performance with the short flankers (Figure 2a and b) but fail to explain the change with long flankers because long flankers increase the amount of irrelevant information (Figure 2a and c). In a second experiment, we showed that a Good Gestalt plays a crucial role in crowding. When the vernier was flanked by two single lines, crowding increased compared to the unflanked condition (Figure 2d). When the two lines were extended to a rectangle,

Citation: Manassi, M., Sayim, B., & Herzog, M. H. (2013). When crowding of crowding leads to uncrowding. *Journal of Vision*, *13*(13):10, 1–10, http://www.journalofvision.org/content/13/13/10, doi:10.1167/13.13.10.

doi: 10.1167/13.13.10 Received March 11, 2013; published November 8, 2013 ISSN 1534-7362

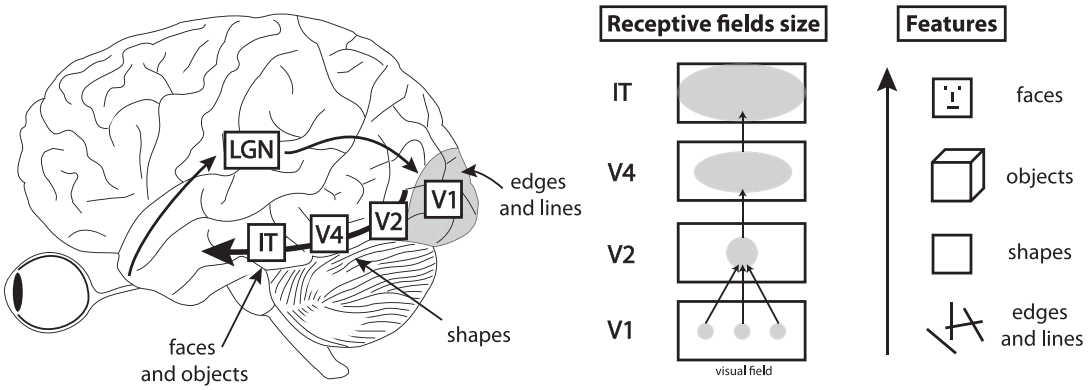

Figure 1. Hierarchical, feedforward visual processing. Stimuli are processed in a series of visual areas. V1 neurons are most sensitive to low-level features, such as edges and lines. In higher visual areas, like V4 and IT, receptive fields are larger, and neurons are sensitive to complex features, such as shapes and objects. Responses of high-level neurons are fully determined by the neural firing of lower-level neurons. For example, the neural firing to a square is determined by the neural firing for two vertical and two horizontal lines.

crowding decreased (Figure 2d and e). Keeping constant the "low-level energy" of the stimulus, crossing the horizontal lines of the rectangles increased crowding compared to the previous condition (Figure 2e and f). Hence, figural aspects strongly matter in crowding (see also Livne & Sagi, 2007). Also increasing flanker size can improve performance (Levi & Carney, 2009; Saarela, Sayim, Westheimer, & Herzog, 2009). Very similar results were also found in foveal vision (Malania, Herzog, & Westheimer, 2007; Sayim, Westheimer, & Herzog, 2008, 2010).

We proposed that crowding is strongly determined by grouping. When the vernier is grouped with the flankers, thresholds increase, and crowding is strong (Figure 2a, d, and f). When the vernier is perceived as standing out from the flankers' configuration, thresholds decrease and crowding is weak (Figure 2b, c, and e). Subjective ratings further supported our hypothesis. For example, we asked subjects to rate how much the

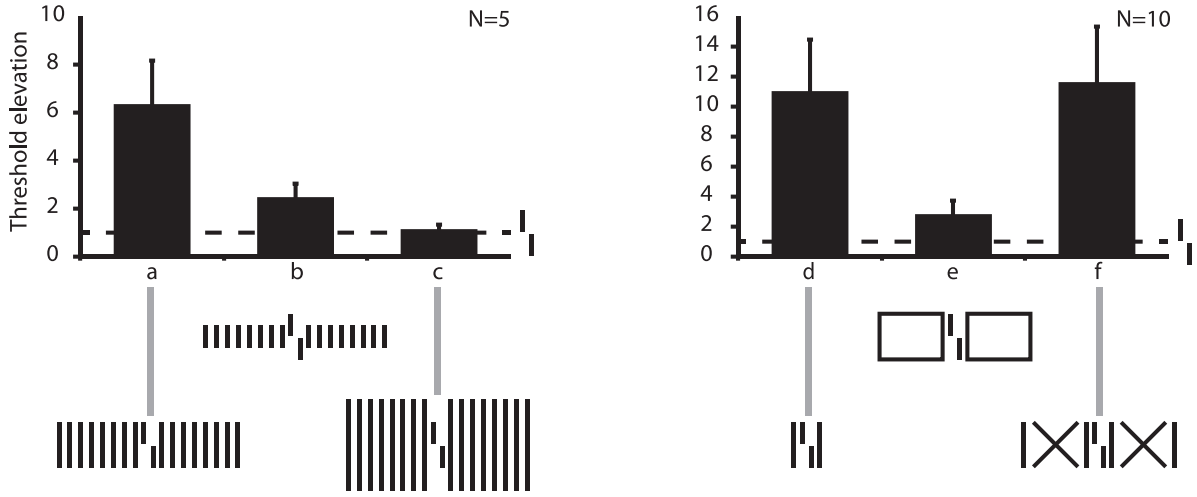

Figure 2. Crowding and grouping (data replotted from Manassi et al., 2012). Stimuli were presented at 3.88° of eccentricity (a through f). Observers indicated whether a vernier was offset to the left or to the right. We determined the offset size for which 75% correct responses occurred (threshold). Results are plotted in terms of threshold elevation compared to a single vernier condition without flankers; i.e., thresholds in the flanking conditions are divided by the threshold of the unflanked condition (dashed lines). A vernier flanked by eight lines of the same length on each side yields a strong threshold elevation (a) compared to the unflanked condition. When the vernier is flanked by eight shorter lines, performance improves (b), and crowding is almost absent for long lines (c). A vernier flanked by two lines of the same length yields a strong threshold elevation (d) compared to the unflanked threshold. When the vernier is flanked by rectangles, performance improves (e). When crossing the horizontal lines of the rectangles, performance deteriorates compared to the previous condition (f).

vernier stands out from the flankers, and the results showed a good qualitative match to the psychophysical results (see Malania et al., 2007; Manassi et al., 2012).

The mechanisms of grouping in crowding are largely unknown. It may, for example, be the case that first textures (arrays of long flankers) and figures (squares) are processed, and crowding occurs subsequently by pooling or substitution within the textures and figures. Here, we show a more complex case, in which crowding on a figural level (between shapes) determines crowding on the more basic level (vernier and square). Contrary to classical models of object recognition, we propose that figural processing determines low-level processing.

Materials and methods

Apparatus

In four experiments, stimuli were presented on a Philips 201B4 CRT monitor driven by a standard accelerated graphics card. Screen resolution was set to 1024 by 768 pixels at a 100-Hz 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. Target and flankers consisted of white lines presented on a black background. The luminance of the stimuli was 80 cd/m^2. Viewing distance was 75 cm.

Stimuli and task

In the square/diamond task (Figure 3 upper panel, Figure 6), observers were asked to indicate whether the central rectangle/diamond was wider along the horizontal or the vertical axis. In the vernier task (Figure 3 lower panel, Figure 4-5), observers were asked to indicate the offset direction of the vernier. The vernier stimulus consisted of two vertical 40-arcmin-long lines separated by a vertical gap of 4 arcmin. Left and right offsets were balanced within each block. Squares and diamonds were made up of four lines that were 2° long. Flanker configurations were centered on the central shape/vernier and were symmetrical in the horizontal dimension. The distance between each flanking shape was 0.5°. Each configuration was presented at an eccentricity of 9° to the right of a fixation point. Eccentricity refers to the center of the target location (shape or vernier). To reduce target position uncertainty, we added two vertical lines (40-arcmin long), 150 arcmin above and below the center of the target. Target and flankers were presented simultaneously for 150 ms. The starting width of the shape was 16.66 arcmin. The starting vernier offset was 16.66 arcmin.

In Experiment 3, we increased the number of squares. To obtain a strong crowding effect, we adjusted the square size and the spacing between squares for each observer individually. We started with a square size of 2° and an intersquare spacing of 0.5°. If the vernier offset threshold was not at least seven times the threshold of the unflanked vernier, we reduced square size and spacing by 5%. Four observers showed strong interference from the beginning; for four observers, we used 80% and for two observers 75%.

Procedure

Observers were instructed to fixate on the fixation point during the trial. After each response, the screen remained blank for a maximum period of 3 s, during which the observer was required to make a response by pressing one of two push buttons. The screen was blank for 500 ms between each response and the next trial. An adaptive staircase procedure (Taylor & Creelman, 1967) was used to determine the square width or vernier offset for which observers reached 75% correct responses. Thresholds were determined after fitting a cumulative Gaussian to the data using probit and likelihood analyses. In order to avoid extremely large vernier offsets, we restricted the PEST procedure to 33.32 arcmin (i.e., twice the starting value). 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. Auditory feedback was provided after incorrect or omitted responses.

Statistics

Threshold data were analyzed with a repeated measures ANOVA. Tukey's post-hoc tests were used for pairwise comparisons for all flanker configurations. All comparisons reported in the Results section were significant.

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

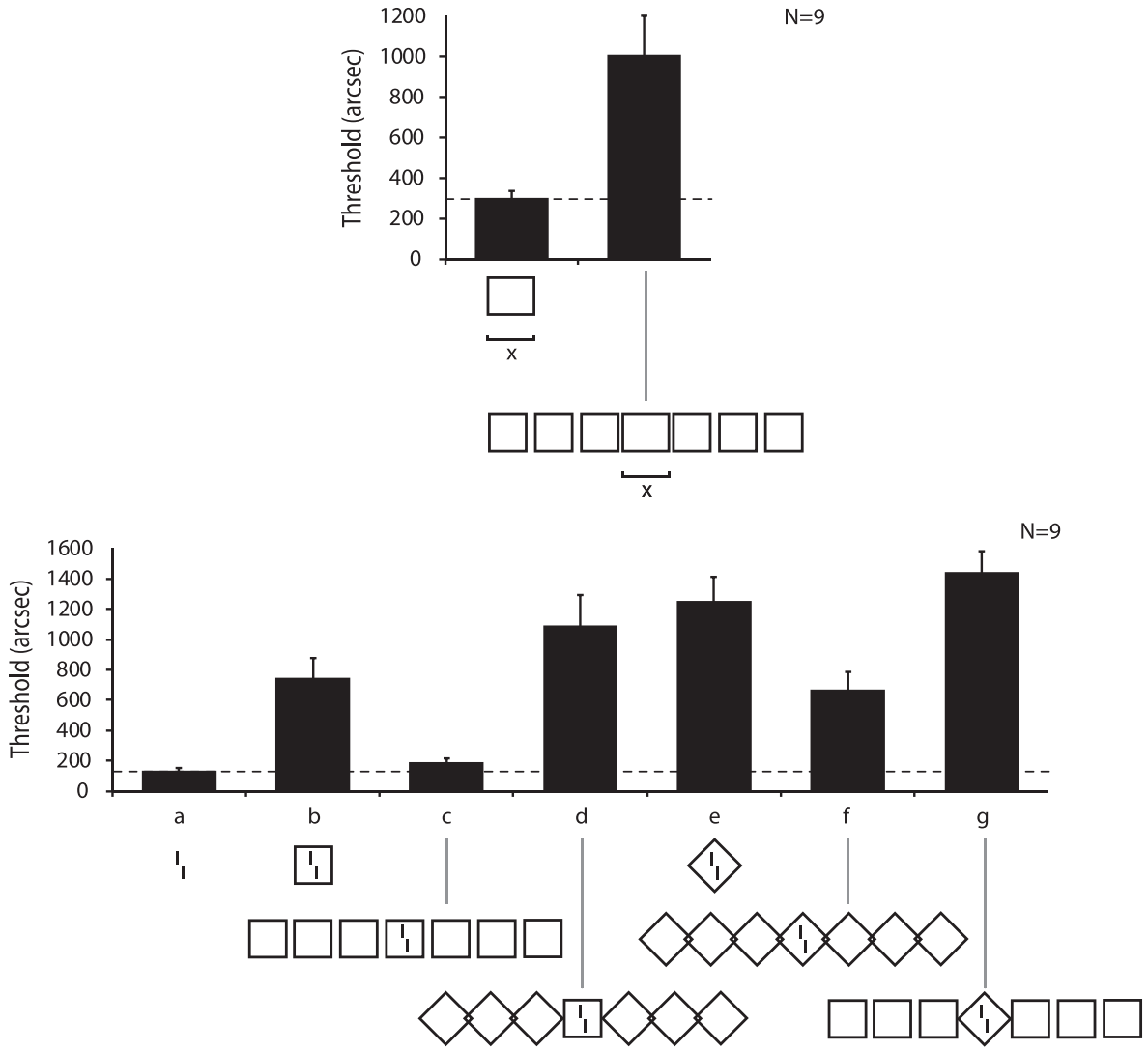

Figure 3. Upper panel: Observers were asked to discriminate whether a rectangle was wider along the horizontal or vertical axis ("x"). We determined the threshold width for which 75% correct responses were obtained. When the rectangle was flanked by three squares on each side, thresholds strongly increased compared to when presented alone. This is a classic crowding effect. Lower panel: (a) Observers were asked to discriminate the vernier offset direction (left vs. right). We determined the offset size for which 75% correct responses occurred (dashed line). (b) Thresholds increased when the vernier was embedded in a square. (c) Compared to the single-square condition, thresholds decreased and crowding almost vanished when the square was flanked by additional squares. Hence, crowding of crowding led to *un*crowding. We propose that inhibition between the squares led to weaker neural representations of the central square and, thus, to less inhibition on the vernier. (d) Thresholds increased when rotating the flanking squares by 45°, creating six diamonds. We suggest that, because of the different shapes, the central square and the neighboring diamond representations do not inhibit each other. (e through g) Thresholds increased compared to the unflanked condition when the vernier was embedded in a diamond (e). Thresholds decreased when the diamond was flanked by other diamonds (f). Thresholds increased when rotating the neighboring diamonds by 45° (g). Error bars indicate the standard error of the mean of nine observers.

Test (Bach, 1996). Observers were told that they could quit the experiment at any time they wished. Nine observers (four females) performed Experiment 1 (Figure 3), and 10 observers (four females) performed Experiment 2 (Figure 5). Ten observers (four females) participated in Experiment 3 (Figure 4), and five observers (one female) participated in Experiment 4 (Figure 6).

Results

Experiment 1

We presented a rectangle at 9° of eccentricity. Observers were asked to discriminate whether the rectangle was wider along the horizontal or vertical axis

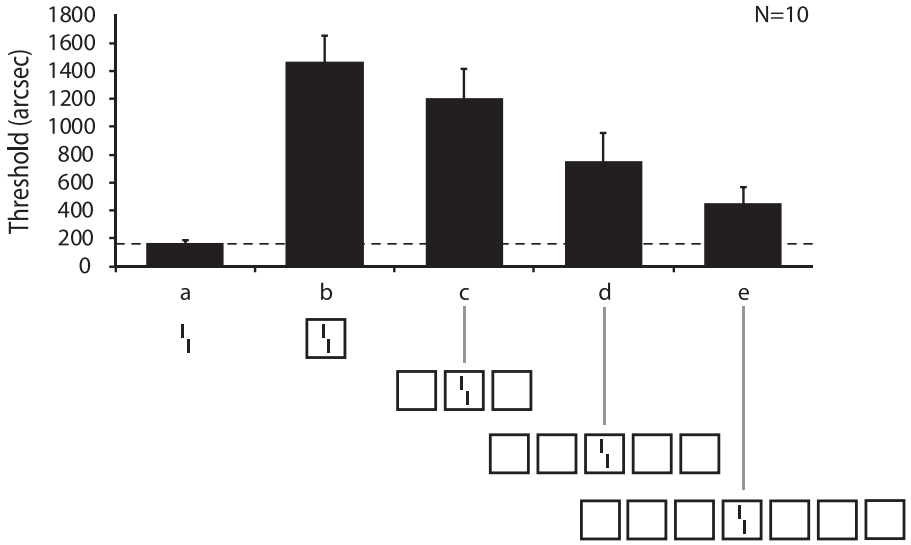

Figure 4. (a) Observers were asked to discriminate the vernier offset direction (left vs. right). (b) Thresholds increased when the vernier was embedded in a square. (c through e) Thresholds gradually decreased when increasing the number of flanking squares.

(Figure 3, upper panel). Thresholds were around 300 arcseconds. Thresholds strongly increased when the rectangle was flanked by six squares (Figure 3, upper panel). This is a classic crowding effect. Next, we presented a vernier within the square and asked observers to discriminate the vernier offset direction (left or right). Threshold increased compared to when the vernier was presented alone—another classic crowding effect (Figure 3a and b, lower panel). We then combined the two conditions; i.e., we presented the vernier within a central square that was flanked by three squares on each side. Classic models of crowding predict that the vernier should be severely crowded because it is, first, crowded by the square and, second, by the other squares. However, the opposite is the case. Crowding almost completely vanished; i.e., perfor-

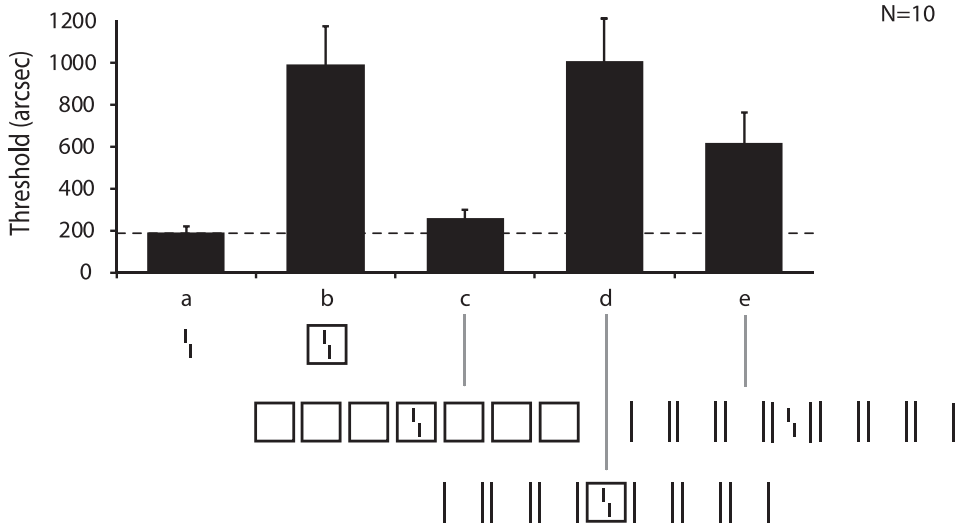

Figure 5. For 10 new observers, thresholds increased, as in the last experiment, when the vernier was embedded in a square (a and b). Thresholds decreased when six additional squares were presented (b and c). When only the vertical lines of the neighbouring squares were presented, thresholds were on the same level as with one square (b compared to d). Hence, the flanking vertical lines themselves do not lead to uncrowding. Also, when only vertical lines were presented, thresholds remained on a high level (e). Hence, interactions between the flanking lines of the central square and the flanking lines of the neighboring squares cannot explain *un*crowding.

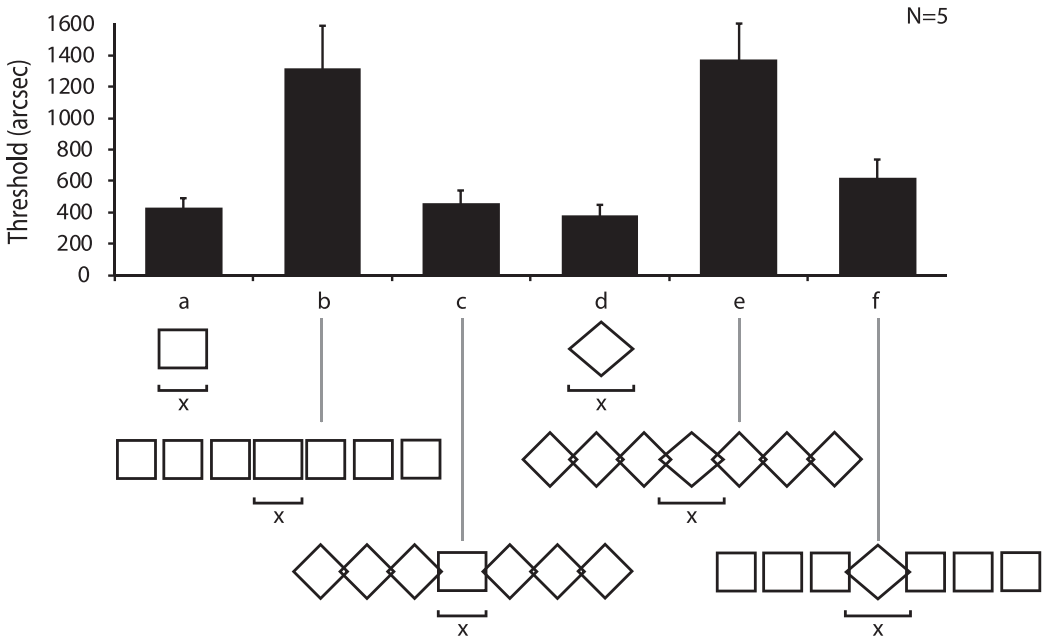

Figure 6. Observers were asked to discriminate whether a rectangle (a through c) or a diamond (d through f) was wider along the horizontal or vertical axis. (b) Thresholds increased compared to the single-rectangle condition when the rectangle was flanked by other six squares. (c) Thresholds decreased when rotating the flanking squares by 45°, creating six diamonds. (d through f) Same pattern in inverse conditions. Observers were asked to discriminate whether the central diamond (d) was wider along the horizontal or vertical axis. (e) Thresholds increased compared to the single-diamond condition in which the diamonds were flanked by six other diamonds. (f) Thresholds decreased when rotating the neighboring diamonds by 45°, creating six squares.

mance of the "doubly crowded" vernier was almost as good as when presented alone (compare Figure 3b and c). Crowding of crowding led to *un*crowding.

We propose that the uncrowding effect can be best explained by figural interactions. In the multisquare condition (Figure 3c), the seven squares crowd each other because of figural identity (shape similarity). Because of this crowding, the vernier representation is little or not affected by crowding of the central square and, thus, crowding diminishes.

According to this hypothesis, we expect strong crowding when figural similarity changes. To test this prediction, we rotated the neighboring squares by 45°, creating six diamonds (Figure 3d). As expected, thresholds strongly increased. We suggest there is weaker inhibition between squares and diamonds because of their dissimilar shapes even though low-level properties, such as pixel energy, are the same. The central square representation is not inhibited and, thus, inhibition on the vernier is strong (Figure 3d). Similar effects were observed when the vernier was embedded in a diamond. Thresholds increased compared to the vernier-alone condition (compare Figure 3a and e). When the vernier embedded in the diamond was flanked by three other diamonds on each side, thresholds decreased (compare Figure 3e and f). When the neighboring diamonds were rotated by 45°, creating six squares, thresholds increased compared to the condition in which only diamonds were presented (compare Figure 3f and g).

Experiment 2

Next, we show that "uncrowding" occurs gradually, depending on the number of squares (Figure 4). As before, we embedded the vernier in a square, and thresholds increased (Figure 4b). We then increased the number of flanking squares from one to seven. Threshold gradually decreased as more squares were presented (Figure 4c through e).

Experiment 3

Can low-level interactions between the lines making up the squares and the vernier explain the "uncrowding" effect? We first repeated the two main conditions for 10 new observers (Figure 5). As before, thresholds increased when the vernier was embedded in a square, and thresholds decreased when adding six further squares (compare Figure 5a and c). Next, we presented only the vertical lines making up the flanking squares. Still, crowding was strong, with thresholds as high as in the single-square condition (compare Figure 5b and d).

Similarly, by presenting only the vertical lines of *all* squares, thresholds remained high, excluding the possibility that line–line rather than square–square interactions explained the *un*crowding effect (compare Figure 5c and e).

Experiment 4

Finally, we show that vernier crowding and crowding of the central shape show similar characteristics. As in the first experiment, observers indicated whether the central square was wider along the horizontal or vertical axis (Figure 6). Thresholds were around 400 arcseconds (Figure 6a). When the rectangle was flanked by 2 × 3 squares, thresholds increased compared to the single rectangle condition (compare Figure 6a and b). When rotating the flanking squares by 45°, creating six diamonds, thresholds decreased, and crowding ceased almost completely (compare Figure 6b and c). We found a similar pattern for the inverse conditions. Observers indicated whether the diamond was wider along the horizontal or vertical axis. For one diamond, thresholds were on the same level as in the single-rectangle condition (compare Figure 6a and d). When the diamond was flanked by 2 × 3 diamonds, thresholds strongly increased compared to the single-diamond condition (compare Figure 6d and e). When rotating the flanking diamonds by 45°, creating six squares, thresholds strongly decreased compared to the condition in which only diamonds were presented (compare Figure 6e and f). Uncrowding is best explained by the differences in shape (Figure 6c and f; Kooi, Toet, Tripathy, & Levi, 1994).

Discussion

Crowding occurs when a target is neighbored by flankers. Most models explain crowding by local interactions with only nearby flankers deteriorating performance. For this reason, most crowding experiments have used only single flankers next to the target (e.g., Levi, Hariharan, & Klein, 2002; Levi, Klein, & Hariharan, 2002; Pelli et al., 2004; Strasburger et al., 1991; Toet & Levi, 1992). Bouma's law proposes that crowding occurs only in a window with a size of about half of the eccentricity of target presentation, which is roughly in accordance with the cortical magnification factor (Bouma, 1970; Pelli et al., 2004; Pelli & Tillman, 2008). Within this window, features of the target and the flankers are thought to be pooled (Parkes et al., 2001; Wilkinson et al., 1997), substituted (Huckauf & Heller, 2002; Strasburger et al., 1991), or impossible to "access" by attention (He et al., 1996). All models predict that crowding increases when the number of flankers increases within the critical window because more irrelevant features are pooled or are substituted. Flankers outside Bouma's window do not affect performance.

However, increasing the number or the size of flankers can, surprisingly, strongly improve performance (Levi & Carney, 2009; Malania et al., 2007; Manassi et al., 2012; Saarela et al., 2009; Sayim et al., 2010; Wolford & Chambers, 1983; see also Figure 3). Performance can even approach the performance level in the unflanked, target-alone condition (Figure 3c). These results cannot easily be explained by "local models." We propose that grouping is *necessary* in crowding. Without grouping of target and flankers, there is no crowding (except for strong differences between target and flankers in contrast or luminance). However, grouping is not sufficient for crowding. For example, when target and flankers are spatially very distant, target and flankers may still group, but there is no crowding. Grouping can occur by many cues as the Gestaltists have shown. For example, the long flankers in Figure 2 group by similarity, making up two textures to the left and right of the vernier. Because of the size difference, the vernier is not grouped with the textures and maintains its "identity." It is an independent object. In this condition, performance is far superior compared to the condition in which flankers have the same length as the vernier and all elements make up one texture. Single-line flankers can lose their "crowding power" also when becoming part of good Gestalts, such as the squares in Figure 2e. The rationale is the same as before. Because of the different Gestalts, the vernier is not grouped with the squares, and, hence, crowding is weak compared to when the single flankers are presented in isolation. Grouping plays a key role also in the current experiments. We suggest that the vernier groups with the central square because of the Gestalt principle of common region (Palmer, 1992). When the central square is flanked by neighboring squares, all squares group by shape similarity, which releases the grouping by common region, leading to ungrouping of the vernier from the central square. We like to mention that, as with many other paradigms, it is not easy to predict performance and crowding strength when Gestalt cues are combined.

Even though target-flanker grouping is crucial in crowding, grouping does not explain why performance deteriorates. Grouping is "neutral" to this question. An inhibitory or otherwise "suppressive" mechanism is needed to explain the deleterious effects of crowding. One option is that, *within* groups, elements are pooled, substituted, or cannot be "accessed" by attention. Other mechanisms might be related to pooling by summary statistics (Balas, Nakano, & Rosenholtz, 2009).

Whereas the current experiments do not address the issue of "suppressive mechanisms", they provide new insights about the interactive processing of target and flankers. Our results show that different types of crowding, related to different levels of processing, can strongly interact with each other. When the central square was flanked by neighboring squares, vernier offset discrimination strongly improved compared to the single-square condition (Figure 3b and c). With flanking diamonds, there was strong crowding. Hence, "global" crowding within the multisquare array strongly interacted with the rather local central-square–vernier crowding. Crowding–crowding interactions are only one potential mechanism. Most likely, in our previous experiments, the long lines did not crowd each other strongly, still leading to ungrouping of the vernier (Figure 3c). Hence, our results show that crowding can occur on many levels and that these levels can mutually interact with each other in agreement with previous findings showing that crowding can occur on many levels in the visual hierarchy (Louie, Bressler, & Whitney, 2007; Wallace & Tjan, 2011; Whitney & Levi, 2011). Our results add to these results that, surprisingly, higher-level crowding can undo lower-level crowding.

Our *un*crowding effects occur in a much larger window than predicted by Bouma's law (Bouma, 1970). In Figure 4, the vernier target was presented at 9° of eccentricity. Hence, Bouma's window is $9/2 = 4.5°$. Still, crowding strongly decreased when the third and fourth outer squares were added at 5° and 7.5° from the target, respectively.

Our results have strong implications for vision in general. In most simple models, visual information processing proceeds from low-level (edges and lines) to high-level (objects and shapes) analysis. If information is lost at the early stages, it is irretrievably lost. However, it seems that, in our experiments, the constituting lines of the squares are processed first. However, target vernier processing is not affected at this stage. Second, the squares are computed from the lines. Third, the square representations inhibit each other. This inhibition occurs on a figural level, i.e., between neurons sensitive to squares, similarly to neurons, tuned to the same orientation at the early stages of vision. When the central and flanking shapes are dissimilar (central square and flanking diamonds; Figure 3d), there is no or only weak inhibition. Finally, when the central square is inhibited, the vernier is disinhibited (multisquare condition vs. flanking diamonds condition, Figure 3c and d). Hence, inhibition between the squares leads to disinhibition of the vernier. This scenario is supported by the fact that performance improved as more squares were added (Figure 4) and that different flanking shapes (squares vs. diamonds) did not lead to uncrowding of the vernier even though pixel energy was similar.

In an alternative scenario, the whole stimulus configuration is analyzed first, and different objects (or groups) are computed without any inhibition mechanism involved. In the multisquare condition, the vernier stands out from the structure of identical shapes and may be processed potentially in an independent "channel" (Figure 3c and f). In the crowded conditions, the vernier is part of one structure together with the central square, processed within one channel (Figure 3d and g).

Regardless of which scenario is true, the strong effects of uncrowding question the intuitive idea that there is a simple link between basic neural processes, such as pooling of nearby neural signals, and perception. The human brain does not stereotypically integrate nearby features across the visual field. There is the intermediate step of grouping. Hence, before we can understand crowding, we need to understand how elements group. In this sense, we propose that grouping *precedes* pooling (Parkes et al., 2001; Wilkinson et al., 1997), substitution (Huckauf & Heller, 2002; Strasburger et al., 1991), inhibition, or any suppression mechanism involved in crowding. Unfortunately, grouping is not a well-defined issue, and its mechanisms remain elusive since the days of the Gestaltists.

Our results are in agreement with other paradigms in which wholes determine the appearance of parts (e.g., Hochstein & Ahissar, 2002; Pomerantz & Portillo, 2011; Weisstein & Harris, 1974; Wertheimer, 1923). We found that larger masks improve performance also in pattern and metacontrast masking (Duangudom, Francis, & Herzog, 2007; Herzog & Fahle, 2002). In surround suppression, contrast discrimination of Gabors improved when the target Gabor ungrouped from the surround (Saarela & Herzog, 2009). Hence, grouping seems to be a crucial, intermediate process in many visual paradigms.

Keywords: crowding, grouping, contextual modulation, vernier acuity, hierarchical processing

Acknowledgments

We thank Marc Repnow for technical support and Aaron Clarke for useful comments on the manuscript. This work was supported by the Project "Basics of visual processing: what crowds in crowding?" of the Swiss National Science Foundation (SNF). Bilge Sayim is currently funded by an FWO Pegasus Marie-Curie grant.

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), Lausanne, Switzerland.

References

Bach, M. (1996). The Freiburg visual acuity test–Automatic measurement of visual acuity. *Optometry and Vision Science, 73*(1), 49–53. http://www.hubmed.org/display.cgi?uids=8867682.

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision, 9*(12):13, 1–18, http://www.journalofvision.org/content/9/12/13, doi:10.1167/9.12.13. [PubMed] [Article]

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177–178. http://www.ncbi.nlm.nih.gov/pubmed/5437004.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron, 73*(3), 415–434. Available from http://www.hubmed.org/display.cgi?uids=22325196, doi:10.1016/j.neuron.2012.01.010.

Duangudom, V., Francis, G., & Herzog, M. H. (2007). What is the strength of a mask in visual metacontrast masking? *Journal of Vision, 7*(1):7, 1–10, http://www.journalofvision.org/content/7/1/7, doi:10.1167/7.1.7. [PubMed] [Article]

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding changes appearance. *Current Biology*, *20*(6), 496–501. http://www.ncbi.nlm.nih.gov/pubmed/20206527.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334–337. http://www.hubmed.org/display.cgi?uids=8848045.

Herzog, M. H., & Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature, 415*(6870), 433–436. http://www.hubmed.org/display.cgi?uids=11807555.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36*(5), 791–804. http://www.ncbi.nlm.nih.gov/pubmed/12467584.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148,* 574–591. http://www.ncbi.nlm.nih.gov/pubmed/14403679.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160,* 106–154. http://www.ncbi.nlm.nih.gov/pubmed/14449617.

Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. *Visual Cognition, 9*(2), 889–910.

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. http://www.ncbi.nlm.nih.gov/pubmed/7993878.

Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology, 19*(23), 1988–1993. http://www.sciencedirect.com/science/article/pii/S0960982209017692.

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):1, 1–12, http://www.journalofvision.org/content/7/2/1, doi:10.1167/7.2.1. [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]

Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision, 12*(10):13, http://www.journalofvision.org/content/12/10/13, doi:10.1167/12.10.13. [PubMed] [Article]

Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology*,

24(3), 436–447. http://www.ncbi.nlm.nih.gov/pubmed/1516361

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. http://www.ncbi.nlm.nih.gov/pubmed/11426231.

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. http://www.ncbi.nlm.nih.gov/pubmed/18828191.

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: Toward a theory of basic gestalts. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1331–1349. http://www.ncbi.nlm.nih.gov/pubmed/21728463.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. http://www.ncbi.nlm.nih.gov/pubmed/10526343.

Saarela, T. P., & Herzog, M. H. (2009). Size tuning and contextual modulation of backward contrast masking. *Journal of Vision, 9*(11):21, 1–12, http://www.journalofvision.org/9/11/21, doi:10.1167/9.11.21. [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]

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. http://www.ncbi.nlm.nih.gov/pubmed/20483840

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.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41*(4), 782–787.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. http://www.ncbi.nlm.nih.gov/pubmed/8632824.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32*(7), 1349–1357. http://www.hubmed.org/display.cgi?uids=1455707.

Wallace, J. M., & Tjan, B. S. (2011). Object crowding. *Journal of Vision, 11*(6):19, 1–117, http://www.journalofvision.org/content/11/6/19, doi:10.1167/11.6.19. [PubMed] [Article]

Weisstein, N., & Harris, C. S. (1974). Visual detection of line segments: An object-superiority effect. *Science, 186*(4165), 752–755. http://www.ncbi.nlm.nih.gov/pubmed/4417613.

Wertheimer, M. (1923). Untersuchungen zur lehre von der gestalt; ii. Psychologische forschung. *Psychologische Forschung,* (*4*), 301–350.

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. http://www.ncbi.nlm.nih.gov/pubmed/21420894.

Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America*, *14*(9), 2057–2068. http://www.ncbi.nlm.nih.gov/pubmed/9291601.

Wolford, G., & Chambers, L. (1983). Lateral masking as a function of spacing. *Perception & Psychophysics, 33*(2), 129–138. http://www.hubmed.org/display.cgi?uids=6844104.