

Release of crowding by pattern completion

Mauro Manassi

Laboratory of Psychophysics, Brain Mind Institute,
École Polytechnique Fédérale de Lausanne,
Lausanne, Switzerland



Frouke Hermens

School of Psychology, University of Aberdeen,
Aberdeen, UK



Gregory Francis

Department of Psychological Sciences,
Purdue University, West Lafayette, IN, USA



Michael H. Herzog

Laboratory of Psychophysics, Brain Mind Institute,
École Polytechnique Fédérale de Lausanne,
Lausanne, Switzerland



In crowding, target perception deteriorates in the presence of flanking elements. Crowding is classically explained by low-level mechanisms such as pooling or feature substitution. However, we have previously shown that perceptual grouping between the target and flankers, rather than low-level mechanisms, determines crowding. There are many grouping cues that can determine crowding, such as low- and high-level feature similarity, low- and high-level pattern regularity, and good Gestalt. Here we show that pattern completion, another grouping cue that is important for crowding in foveal vision, is also important in peripheral vision. We also describe computer simulations that show how pattern completion, and crowding in general, can be partly explained by recurrent processing.

Introduction

In crowding, perception of a target deteriorates in the presence of neighboring elements. For example, it is easy to identify a letter when it is presented alone. However, identification strongly deteriorates when the same letter is flanked by other letters (Andriessen & Bouma, 1976; Levi, 2008; Pelli, Palomares, & Majaj, 2004; Pelli & Tillman, 2008). According to pooling models, crowding occurs when neurons from higher visual areas with larger receptive fields pool information from lower level neurons with smaller receptive fields, which leads to averaging of target and flanker signals (Balas, Nakano, & Rosenholtz, 2009; Dakin,

Cass, Greenwood, & Bex, 2010; Greenwood, Bex, & Dakin, 2009, 2010; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli et al., 2004; Pöder, 2012; van den Berg, Roerdink, & Cornelissen, 2010; Wilkinson, Wilson, & Ellemberg, 1997). According to substitution models, crowding occurs because features of the target are confused with features of the flankers (Hanus & Vul, 2013; Huckauf & Heller, 2002; Strasburger, 2005; Strasburger, Harvey, & Rentschler, 1991). Both basic pooling and substitution models predict that when the number or size of flankers increases, crowding increases because more irrelevant information is pooled or more features are confused.

We have previously shown that this prediction is not always true. We determined offset discrimination thresholds for verniers flanked by various line configurations. When the vernier was flanked by eight flankers on each side with the same length as the vernier, crowding was strong (equal-length condition). When we either increased or decreased the length of the flankers, crowding was much weaker (foveal vision: Malania, Herzog, & Westheimer, 2007; peripheral vision: Manassi, Sayim, & Herzog, 2012). Likewise, crowding strongly decreased when we increased the number of short flankers from 2 to 8 or 16 flankers (Manassi et al., 2012). Very similar results were found with Gabor stimuli (Saarela, Sayim, Westheimer, & Herzog, 2009). We proposed that one of the best predictors of crowding strength is the extent to which target and flanking elements group together. When the target groups with the flankers (equal-length condi-

Citation: Manassi, M., Hermens, F., Francis, G., & Herzog, M. H. (2015). Release of crowding by pattern completion. *Journal of Vision*, 15(8):16, 1–15, doi:10.1167/15.8.16.

tion), crowding is strong. When the flankers group and the target stands out (short- and long-length conditions), crowding is weak (Malania et al., 2007; Manassi et al., 2012; see also Manassi et al., 2013; Saarela et al., 2009; Saarela, Westheimer, & Herzog, 2010; Sayim, Westheimer, & Herzog, 2008, 2010; Wolford & Chambers, 1983).

Grouping can determine crowding by many cues. First, grouping can occur by low-level similarity such as line length, as shown in the example above. Second, grouping can occur when single elements become part of a good Gestalt. As before, we presented a vernier flanked by two same-length lines. As expected, offset discrimination thresholds strongly increased compared to the unflanked condition. However, when the same two lines became part of flanking rectangles, crowding strongly decreased compared to the two-lines condition. Hence, good Gestalt plays a crucial role in crowding (foveal vision: Sayim et al., 2010; peripheral vision: Manassi et al., 2012).

Third, spacing regularity. Saarela et al. (2010) presented a letter T in the periphery and asked observers to discriminate its orientation. When the spacing between all letters, both target and flankers, was the same, crowding was strong. When the spacing between flanking letters was different compared to the spacing between the target and its direct neighbors, crowding was weaker. It remains unclear whether spacing regularity can be seen as a Gestalt cue for similarity. Another case is pattern regularity, which taps into higher order structural aspects of the stimulus. When a red vernier was flanked on each side by flankers with the same color, crowding increased compared to the unflanked threshold. When the color of the flankers was changed from red to green, crowding diminished. However, when alternating the color of the flankers in a regular green–red fashion, crowding increased compared to the previous condition (foveal vision: Sayim et al., 2008; peripheral vision: Manassi et al., 2012).

Fourth, figural grouping. When a vernier was embedded in a square, crowding increased compared to the unflanked threshold. When additional squares were presented, crowding almost disappeared. When the flanking squares were rotated by 45°, crowding was strong again (peripheral vision: Manassi et al., 2013).

Fifth, pattern *completion*. In foveal vision, Hermens, Herzog, and Francis (2009) combined crowding with forward and backward masking. First, a vernier was presented alone and observers discriminated its offset. Next, in the basic crowding condition, the vernier was flanked by 12 aligned verniers on each side. As expected, crowding was strong. Next, an aligned vernier, which we will call “the mask”, was presented at the location where the target vernier was presented. The mask was presented at various stimulus onset asynchronies

(SOAs), preceding or following the basic crowding configuration. For long SOAs of –200 or 200 ms, foveal crowding was, not surprisingly, at about the same level as without the mask. However, for shorter SOAs, crowding did not increase as one might have expected from combining two powerful deleterious techniques. Quite to the contrary, crowding strongly decreased. The decrease in crowding was explained by pattern completion. The mask vernier fits perfectly in the gap of the two arrays of flanking verniers, thus making up a regular grating with equally spaced, identical elements. These elements are grouped as one entity, and for this reason, the target vernier is released from crowding. Perceptually, the target vernier appears superimposed on the grating and brighter than the grating elements. This pattern completion effect is reminiscent of the *shine-through* masking effect, where a vernier is presented for a very short time (usually 20–50 ms) and followed by a grating of lines presented for 300 ms. For five grating elements, shine-through is absent and performance is strongly impaired compared to the 25-element condition (Herzog & Fahle, 2002; Herzog & Koch, 2001).

Here we show that similar effects also occur in peripheral vision, where crowding usually is investigated (Experiments 1–4). In addition, we show that pattern completion, similar to the shine-through effect, occurs only with an extended number of flankers (Experiment 5) and with more complex flanker layouts (Experiment 6). Finally, we show that *uncrowding* can be partly explained by a recurrent neural model of the shine-through effect (Francis, 2009). Hence, a model tailored to explain temporal phenomena in masking can also explain spatial processing in crowding, linking two seemingly distinct research areas.

General materials and methods

Observers

Observers were paid students of the École Polytechnique Fédérale de Lausanne. 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 as determined by the Freiburg Visual Acuity Test (Bach, 1996). Before the experiments, observers were informed about the general purpose of the experiment and gave their written consent. They were told that they could quit the experiment at any time.

Apparatus and stimuli

Stimuli were presented on an analog HP-1332A XY-display equipped with a P11 phosphor (screen size 12 ×

9.5 cm). The display inputs were controlled by a PC via a custom-made 16-bit DA interface, providing a pixel address resolution of about 1.8 μm . Background luminance of the screen was below 1 cd/m^2 . Luminance of stimuli was 80 cd/m^2 (as measured by a Minolta LS-100 luminance meter). The experimental room was dimly illuminated (0.5 lx), and viewing distance was 75 cm.

We determined offset discrimination thresholds for a vertical vernier. Compared to the stimuli presented by Hermens et al. (2009), we increased the size of the vernier by a factor of 4 to adapt for peripheral vision. The vernier was composed of two vertical lines 40' (arcmin) long separated by a vertical gap of 4'. The starting offset was increased from 1.25' (foveal presentation) to 16.66' (peripheral presentation). The vernier was presented at an eccentricity of 3.88° to the right of a fixation cross (6' diameter). The vernier was presented alone or neighbored by various flanker configurations, which consisted of arrays of aligned verniers. The number of flanking lines was decreased from 24 to 16. We increased the horizontal distance between the vernier and the directly neighboring lines from 3' to 23.33'. Interflanker spacing was also increased from 3' to 23.33'. Vernier and flanker durations were 20 ms.

In addition to the flankers, an aligned vernier (the mask) could be presented at different SOAs. The mask was presented at the same location as the target vernier. In Experiments 5 and 6, the SOA was always 0. Mask duration was 20 ms.

Procedure

Observers were instructed to fixate a cross during the trial. Previous studies with eye tracking showed that observers are able to maintain fixation during the experiment (Manassi et al., 2012). Hence, we did not record eye movements here. After each stimulus presentation, 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 to indicate the offset direction of the target vernier. The screen was blank for 500 ms between the response and the next trial.

An adaptive staircase procedure (Taylor & Creelman, 1967) was used to determine the vernier offset for which observers reached 75% correct responses. Instead of taking the last value of the adaptive procedure as the threshold, we estimated both the threshold and slope of the psychometric function (cumulative Gaussian) by means of maximum-likelihood estimation, taking all data points into account (Wichmann & Hill, 2001).

To avoid strong influences of extremely large vernier offsets on the average data, we restricted the adaptive

staircase procedure to not exceed 33.32', i.e., twice the starting value of 16.66'. Vernier offset thresholds ranged from 50" (arcsec; unflanked threshold) to 2000" (33.32', the maximum offset size allowed by the procedure). To avoid influence of practice and fatigue on the average data, each condition was presented in two blocks of 80 trials each. In the first run, conditions were randomized for each observer. In the second run, the order of conditions was reversed. Auditory feedback was provided after incorrect or omitted responses.

Statistics

In Experiments 1, 3, and 4, we used paired-samples *t* tests to compare the “vernier + flankers” conditions and the “vernier + flankers + mask” conditions for the SOA of 0 ms only. In Experiment 2, where no flankers were presented, we used a paired-samples *t* test to compare the “vernier alone” and “vernier + mask” conditions. In Experiments 5 and 6, threshold data were analyzed with a repeated-measures ANOVA. Tukey's post hoc tests were used for pairwise comparisons for all flanker configurations.

Completion: Similar elements

Experiment 1

Five observers (two females) participated in the experiment. First we determined offset discrimination for a vernier presented alone (Figure 1a, dashed line). Second, the vernier was flanked by 16 same-length flankers (Figure 1a, black horizontal line). Third, in addition to vernier and flankers, a same-length line (the mask) was presented for SOAs of −200, −60, −20, 0, 20, 60, and 200 ms relative to the target.

When the vernier was flanked by 16 same-length flankers, thresholds strongly increased compared to the unflanked condition (Figure 1a, black vs. dashed line). This is a classic crowding effect, in line with previous findings (see Malania et al., 2007, for foveal crowding and Manassi et al., 2012, for peripheral crowding). When the mask was presented in addition, the pattern of results strongly changed. For SOAs around 0, thresholds strongly *decreased* compared to the crowding condition, $t(4) = 3.06$, $p = 0.04$ for SOA = 0, compared to the no-mask 16-flanker condition. For longer SOAs (−200 and 200 ms), thresholds were on a level comparable to the basic crowding condition.

These results are challenging for classic crowding models (Balas et al., 2009; Dakin et al., 2010; Freeman

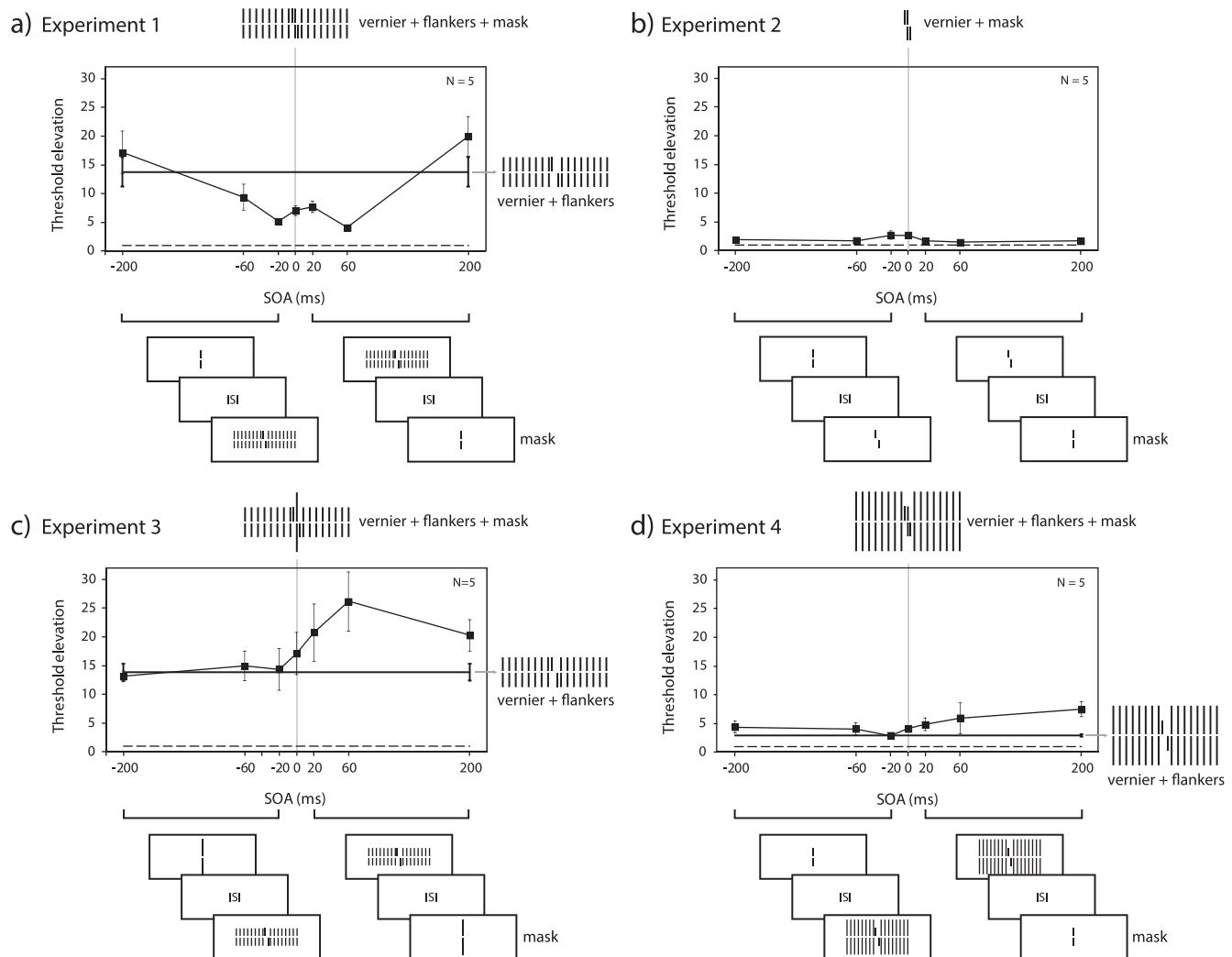


Figure 1. Experiments 1–4. Dashed lines indicate performance in the “vernier alone” condition. The horizontal black lines indicate the basic crowding condition, i.e., vernier with flanking lines (stimulus configuration depicted on the right). In the masking conditions, an aligned vernier was presented at the target location at various SOAs. Negative SOAs indicate that the mask preceded the target, and positive values indicate that the mask followed the target. The vertical gray lines indicate an SOA of 0 ms, i.e., temporal overlap between target, flankers, and mask (stimulus configuration depicted in the upper part). Results are plotted in terms of threshold elevation, i.e., thresholds divided by the threshold of the unflanked condition (dashed lines). A threshold elevation of 1.0 indicates no crowding; values larger than 1.0 indicate crowding. Error bars indicate ± 1 standard error. (a) When the vernier was flanked by the 16 same-length flankers (basic crowding condition), performance strongly deteriorated compared to the “vernier alone” condition (black vs. dashed lines). When the mask was presented in addition, for SOAs of ± 200 ms performance stayed on a constant level compared to the flanking condition. For shorter SOAs, performance strongly improved compared to the flanking condition. (b) When the vernier was preceded or followed by the mask, performance deteriorated very little compared to the unflanked condition. (c) When the vernier was flanked by the 16 same-length flankers, performance strongly deteriorated compared to the “vernier alone” condition (black vs. dashed line). When vernier and flankers were preceded by a central double-length mask (forward masking), performance was on the same level as the flanking condition. For long SOAs (backward masking), performance even deteriorated. (d) When the vernier was flanked by 16 double-length flankers, performance only slightly deteriorated compared to the “vernier alone” condition (black vs. dashed lines). When vernier and flankers were preceded or followed by a mask with the same length as the vernier, performance remained deteriorated.

& Simoncelli, 2011; Greenwood et al., 2009, 2010; Hanus & Vul, 2013; Parkes et al., 2001; Pelli et al., 2004; Pöder, 2012; van den Berg et al., 2010; Wilkinson et al., 1997). According to these models, the target signal is compulsorily pooled or averaged with the

signals of the flankers; hence the combination of crowding and masking can have only detrimental effects. However, we found the opposite result, i.e., masking the crowded vernier led to a decrease in crowding.

How can the *uncrowding-by-masking* results be explained? We propose that when the vernier is flanked by 16 same-length flankers, crowding is strong because of grouping between the vernier and flankers. The mask and flankers together create a grating of regularly aligned verniers for an SOA of 0 ms. This grating is taken as one perceptual identity and thus ungroups from the vernier and reduces crowding. Hence, when the mask “completes” the grating, the vernier is released from crowding strength. This effect is also present for short SOAs of -60 , -20 , 20 , and 60 ms. For long SOAs, crowding reappeared. This interpretation is consistent with observers’ reports: For short SOAs, the vernier is perceived as brighter, as in the shine-through effect of backward masking.

Experiment 2

As a control, we determined vernier offset discrimination thresholds with only the mask—i.e., unlike in Experiment 1, no flankers were presented. Five different observers (two females) performed the experiment. Thresholds increased for short SOAs (-20 , 0 , and 20 ms), but overall the masking effect was very weak (Figure 1b), $t(4) = -4.34$, $p = 0.01$ for SOA = 0 , versus the no-mask condition. Hence, the mask per se has only a very small effect on target discrimination.

Experiments 3 and 4

In Experiment 2, we showed that the mask by itself only weakly affects vernier offset discrimination (Figure 1b). Hence, the decrease in crowding in Figure 1a seems to be due to the completion of the grating by the mask. Here, as further controls, we show that shorter and longer masks, which do not complete the grating pattern, do not lead to a decrease in crowding.

In Experiment 3, we presented the vernier with 16 same-length flankers together with a mask that had twice the length of the other elements (Figure 1c). In Experiment 4, we presented the vernier flanked by 16 double-length flankers with a mask having the same length as the vernier (Figure 1d). Two groups of five observers (one female each) performed the experiments.

When the vernier was flanked by 16 same-length lines, thresholds strongly increased compared to the unflanked condition (Figure 1c, black vs. dashed line). When the double-length mask was presented, thresholds remained on the high level of the basic crowding condition for SOAs of -200 , -60 , -20 , and 0 ms, $t(4) = -1.27$, $p = 0.27$ for SOA = 0 , compared to the no-mask 16-flanker condition. For SOAs of 20 , 60 , and 200 ms, thresholds even increased compared to the crowded condition.

Hence, there was no release from crowding when the mask was longer than the flankers. When the vernier was flanked by 16 double-length lines, thresholds increased compared to the unflanked condition (Figure 1d, black vs. dashed line). In line with previous results (Manassi et al., 2012), long flankers yielded less crowding compared to same-length flankers (Figure 1a, c vs. Figure 1d). When a mask was presented with the same length as the vernier, thresholds remained as high as in the crowded condition—or even increased—for all SOAs, $t(4) = -4.04$, $p = 0.01$ for SOA = 0 , compared to the no-mask condition with 16 long flankers.

We propose that, because of the different line lengths in both experiments, the mask did not group with the flankers and thus the vernier was not released from the interference of the flanking lines. We suggest that only when the mask completes a pattern of *similar* elements is the vernier target released from crowding.

Completion: Number of flankers

In Experiments 1, 3, and 4, we have shown that adding the mask to the flankers strongly increases performance when the mask and flankers make up a regular grating of identical lines. However, there is no decrease of crowding when the lengths of the mask and flankers differ. Hence, it seems that the overall structure of the mask-flanker pattern is crucial for performance improvements. Here we show that indeed, local mechanisms cannot explain improved performance. For example, our results cannot be explained by local interactions between the mask and the flankers directly neighboring the vernier. To this end, we varied the number of flankers and showed that performance improves gradually. A vernier was presented alone or flanked by 2, 4, 8, 12, or 16 same-length lines (Figure 2). In the “mask” conditions, the same-length mask was presented in addition to the flankers. Six observers (three females) performed the experiment. A repeated-measures ANOVA showed a significant interaction between the “mask” condition and the number of flankers, $F(10, 50) = 7.66$, $p < 0.0001$. In the “no mask” conditions, thresholds were strongly elevated compared to the unflanked threshold, irrespective of the number of flankers ($p < 0.05$), meaning that the strength of crowding was independent of the number of flankers. This is consistent with previous reports (Malania et al., 2007; Manassi et al., 2012). In the “mask” conditions with 2, 4, and 8 flankers, there was essentially no difference between the “mask” and “no mask” conditions (strong crowding). For 12 and 16 flankers, however, thresholds strongly decreased compared to the “no mask” condition ($p < 0.05$). The decrease in threshold with 16 flankers in the “mask” condition is

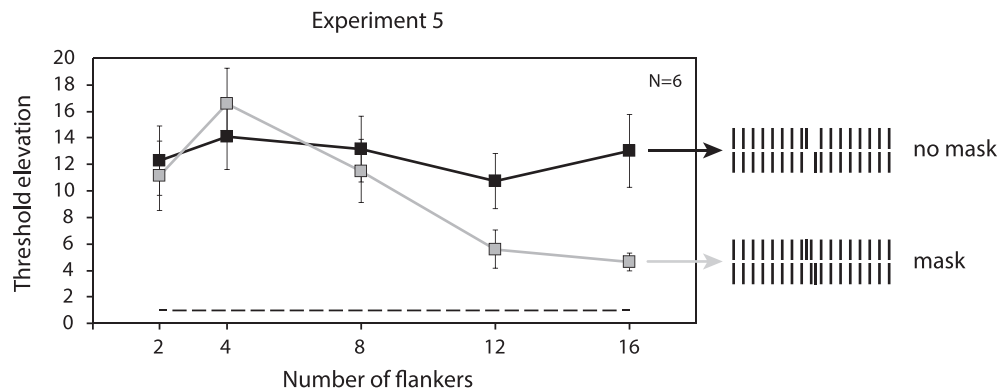


Figure 2. Experiment 5. Threshold elevation as a function of the number of flankers. The dashed line indicates performance of the unflanked vernier. When the vernier was flanked by 2, 4, 8, 12, or 16 same-length flankers, performance stayed at a constant high level. When the same flanker configurations included the mask, performance did not change for 2, 4, or 8 flankers. However, for 12 and 16 same-length flankers, performance improved compared to the “no mask” condition.

consistent with the results reported in Figure 1a (gray vertical line).

Clearly, performance depends on the number of flankers, and hence simple local interactions between the innermost neighbors and the mask cannot explain the improvement of performance, since these innermost flankers are present in all conditions. It remains an open question whether it is the number of flankers or the sheer extension of the flanker array that matters. This question is not easy to answer, because changing the extension of the flanker array while keeping the number of flankers constant also changes the spacing between the flankers (and the vernier) and so makes comparisons impossible.

Completion: Regular patterns

In the previous experiments, we showed that the completion of a pattern of same-length flankers leads to the release of the target vernier from crowding. In Experiment 6, we propose (and show) that the decrease in crowding can be used as a tool to understand the “goodness” of a pattern.

Observers in Experiment 6 were presented with six different flanker configurations. In the first three conditions, we presented eight flankers on each side, spaced at 23.33', as before. The length of flankers gradually decreased from 40' to 5', in steps of 5' (Figure 3a). No mask was used in the first of these conditions. In the second and third conditions, we added a same-length mask (Figure 3b) or a double-length mask (Figure 3c) at the target location. In the next three conditions, the length of flankers was gradually increased from 40' to 75', in steps of 5' (Figure 3d). Again, no mask was used in the first of these three conditions, a same-length mask in the second (Figure

3e), and a double-length mask in the third (Figure 3f). Seven observers (two females) performed the experiment.

The same pattern of results was found for increasing and decreasing flanker lengths, $F(6, 42) = 14.21$, $p < 0.0001$. Without a mask, crowding occurred ($p < 0.05$, Figure 3a, d). When a same-length mask was presented, thresholds strongly decreased compared to the no-mask condition ($p < 0.05$, Figure 3b, e). Double-length masks, in contrast, did not reduce crowding (Figure 3c, f).

We propose that adding the same-length mask creates a regular mask-flankers pattern. Crowding decreases because the vernier ungroups from this pattern (Figure 3b, e). When the longer mask is added, the vernier remains strongly grouped with the flankers because the mask does not make up a regular pattern with the flankers. Crowding remains high (Figure 3c, f).

Neural-network model of perceptual grouping

We have argued that the experimental findings cannot be explained with simple pooling or substitution models because they do not have mechanisms to consider the perceptual grouping of stimulus elements. Similar arguments about perceptual grouping have been proposed in visual masking (shine-through; Herzog & Koch, 2001), where a trailing mask can weaken or enhance the visibility of a target vernier depending on the perceptual grouping of elements in the scene. The standard shine-through effect is produced with stimuli very similar to those used here, although the timing and order are different. Francis (2009) accounted for many properties of the shine-through effects with a neural-network model of visual perception, and the explanations strongly depended on

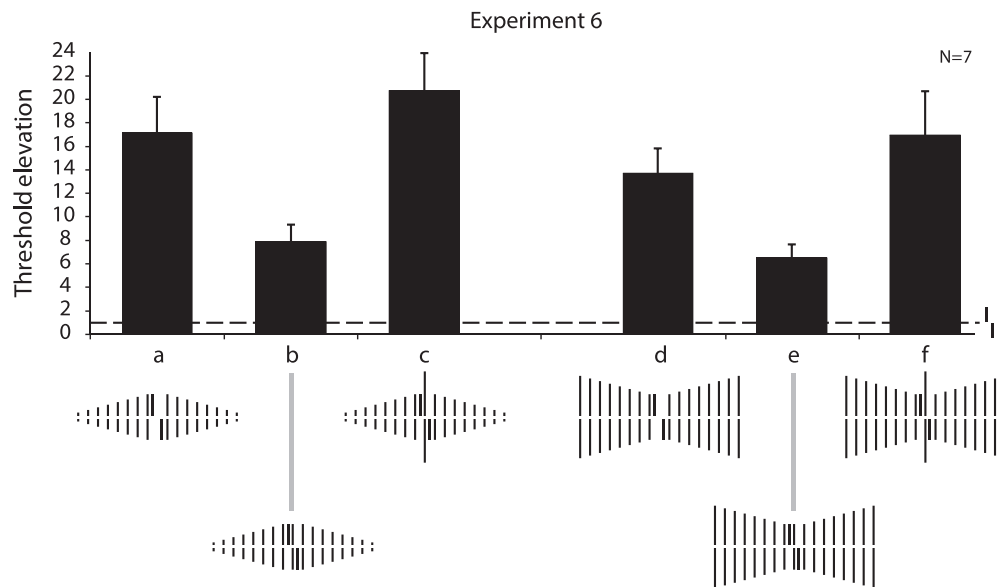


Figure 3. Experiment 6. The dashed line indicates the threshold for the unflanked vernier. (a) Compared to the unflanked condition, vernier offset discrimination deteriorated when the vernier was embedded in a pattern of decreasing-length flankers. (a–b) Performance improved compared to the previous condition when the mask was presented in addition. (b–c) Performance deteriorated when the length of the mask was doubled. (d) Vernier offset discrimination deteriorated when the vernier was embedded in a pattern of increasing-length flankers. (d–e) Performance improved compared to the previous condition when a same-length mask was presented at the target location. (e–f) Performance deteriorated when the length of the mask was doubled.

perceptual grouping of stimulus elements. Given the similarity of the stimuli and task, we wondered if the neural-network model might also explain aspects of crowding that are driven by perceptual grouping. A successful model could provide a definition of grouping that is independent of a measured crowding effect. Such independence is needed in order to be able to test the proposed relation between crowding and grouping.

The model is a dynamic version of the LAMINART model proposed by Cao and Grossberg (2005) to account for stereopsis and 3-D surface perception. On the one hand, the model is quite complicated and contains many characteristics that are not important for an explanation of crowding effects. On the other hand, the model suggests interesting connections between areas of visual perception (such as crowding and stereopsis) that might otherwise seem disparate. Model equations and parameters are given by Francis (2009), and details of the model properties that are unique to the current simulations can be found in the Appendix. The model was not designed to account for crowding data, so, importantly, the parameters are not optimized for the precise conditions of the current experiments. In particular, the simulations were originally designed to account for phenomena that occur in the fovea. Rather than attempt to adjust the parameters to account for various spatial and temporal aspects of peripheral visual processing, we explored the aspects of the model that, based on previous simulations and analysis, should not much depend on the details of

parameter choices. In particular, Francis (2009) showed that the model can produce a shine-through effect by creating an isolated representation of a vernier at a near depth plane. This shine-through vernier is derived from false binocular matches and only forms when the flanking elements group together. This section exposes the model to the stimuli used in Experiments 1–5 and examines whether the model's explanation of the shine-through effect accounts for the observed release from crowding (or its absence).

The model computes template matches for a leftward- and a rightward-shifted vernier and produces a contrast value that corresponds to evidence on whether the image enables discrimination between a left and a right target vernier. Larger discrimination evidence values correspond to better discrimination and lower thresholds, so the vernier discrimination evidence value is plotted in reverse in Figures 4 and 6, which show that the model matches the empirical data reasonably well for Experiments 1, 2, and 5 but does not match the empirical findings for Experiments 3 or 4.

For Experiment 1 (Figure 4a), the findings are explained with the model mechanisms responsible for the appearance of the shine-through effect as described by Francis (2009). Briefly, the shifted elements of the target vernier produce false binocular disparity matches with the mask elements, and these disparity matches can produce a representation of the vernier in a near depth plane when the vernier's representation in a far depth plane is weakened by lateral inhibition and

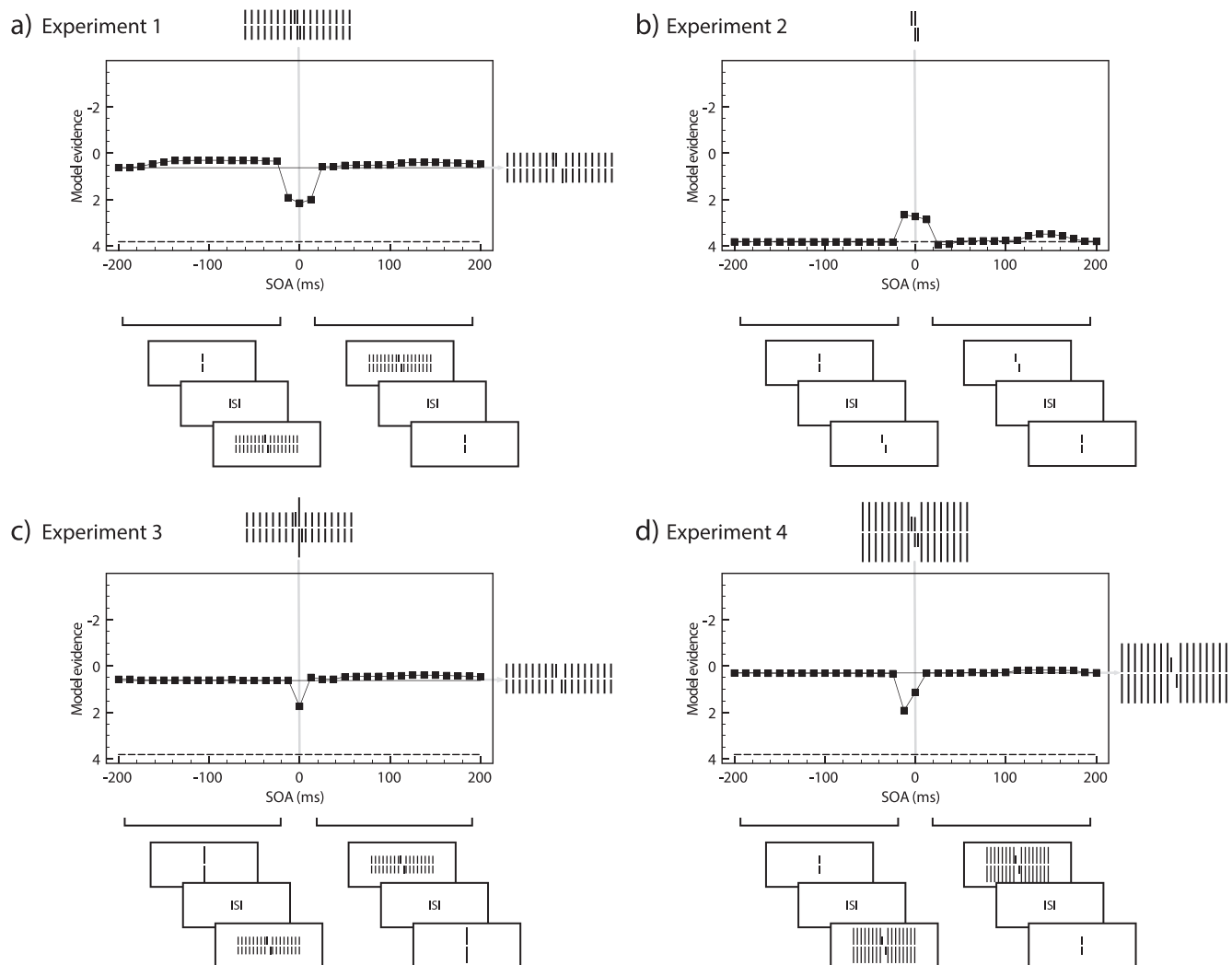


Figure 4. The plots show the model evidence values in reverse order for easy comparison with the empirical thresholds in Figure 1. (a–b) The model captures the main properties of the empirical findings in Experiments 1 and 2. (c–d) The model fails to capture important properties of the findings in Experiments 3 and 4.

perceptual grouping from the flanking elements. Perceptual grouping is indicated in Figure 5a, which shows the activity of orientation-tuned cells shortly after target, mask, and flanker offset for 0 SOA.

Grouping is indicated by the illusory horizontal (black) contours along the top and bottom of the grating elements. This grouping weakens the vertically tuned cells that represent the vernier and thereby disinhibits

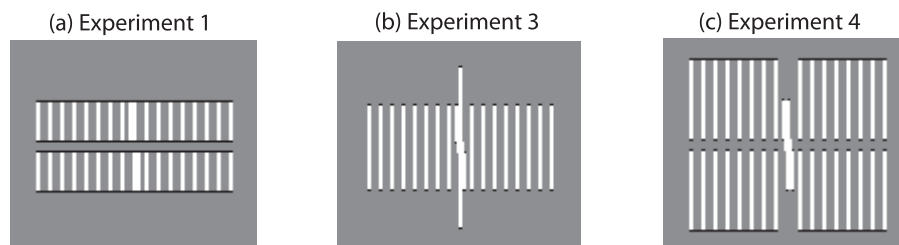


Figure 5. Representations of activity patterns for orientationally tuned cells in the neural-network model. A middle gray pixel indicates no activity, white pixels indicate responses from vertically tuned cells, and black pixels indicate responses from horizontally tuned cells. (a) For the stimuli in Experiment 1, the equal-length flankers group with the mask and target by generating horizontal illusory contours between the elements. (b) Grouping does not occur for the stimuli in Experiment 3 because the long mask blocks the creation of illusory contours between the flanker and target elements. (c) Grouping occurs for the long flankers on each side of the target in Experiment 4, but not across the sides, and the grouping does not include the target.

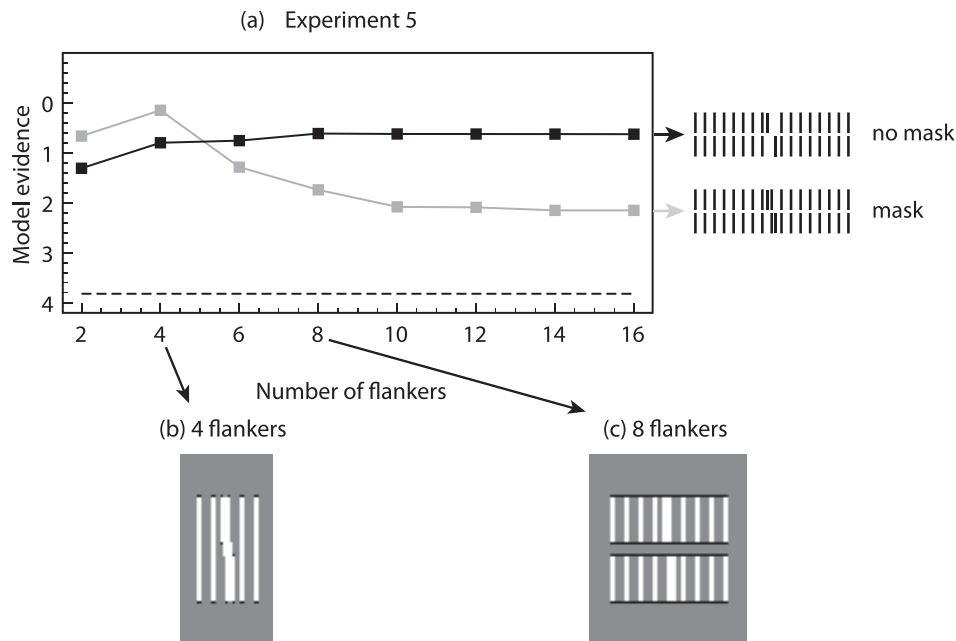


Figure 6. Model evidence related to Experiment 5. (a) Model vernier discrimination is plotted in reverse order for easy comparison with the empirical thresholds in Figure 1. (b) Four flankers are unable to generate the illusory contours that correspond to perceptual grouping. (c) Eight flankers are able to generate illusory contours that correspond to perceptual grouping.

the false binocular matches in a foreground stage (not shown). The disinhibited vernier is represented by itself in the foreground plane, and the template-matching process is then uninfluenced by the mask and flankers (which remain in the background plane). The sensitivity of crowding to the mask SOA reflects the fact that the disparity matches cannot be generated unless the mask and target are in close temporal proximity. When the mask is not present, the false disparity matches are not generated, so there is no shine-through effect, i.e., no release from crowding.

For Experiment 2 (Figure 4b), the weak crowding is due to modest masking from lateral inhibition and the presence of the mask elements in the target templates. The slight increase in crowding for SOAs close to 0 is due to a reverse shine-through effect. Here, the false disparity matches briefly produce a foreground representation of a vernier that is shifted opposite to the actual stimulus (an example of this process is given by Francis, 2009), and since it is isolated from the mask elements, such a representation produces a (brief) strong response in the template. It is interesting that the data in Figure 1b show the same effect. The main property of the simulation is that there is no release from crowding because there are not enough flankers to generate a shine-through effect. This aspect of the model matches the experimental data.

For Experiment 3 (Figure 4c), the model correctly indicates little effect of the mask when it precedes the flankers, but the model does not demonstrate sufficient crowding for positive SOAs. Given that a version of the

model has accounted for a variety of backward masking effects (Francis, 1997), this discrepancy may indicate a need for different simulation parameters. The model also shows improved vernier discrimination (release from crowding) for 0 SOA, but this seems to be for a reason different than in the model's behavior in Experiment 1, where a shine-through effect occurs. In the model simulation of Experiment 3, the neural responses to the target, flankers, and mask representations do not generate shine-through because the long mask prohibits the flanker elements from grouping with the target (note the absence of horizontal illusory contours in Figure 5b). However, after stimulus offset, the representations of these elements deteriorate in such a way that the mask disappears first and thereby partially frees the target from some crowding. Unlike the shine-through effect, which is a robust characteristic of the model, the release from crowding at 0 SOA in Figure 4c may disappear with different stimuli or parameters.

Figure 4d shows that the model behavior does not match the empirical data for Experiment 4, which demonstrates very little crowding for long flankers. Contrary to the empirical findings, the model shows strong crowding for all mask SOAs. Also contrary to the empirical findings, the addition of the mask does not increase crowding and even provides release from crowding for SOAs close to 0. This release is the result of a shine-through effect that generates an isolated vernier representation in the near depth plane. However, for all other SOAs, the signals corresponding to

the target vernier are in the far depth plane with the mask and flanker signals; and since all the signals are in the same depth plane, the mask and flanker elements interfere with the template matches and thereby produce substantial crowding. Another discrepancy is that the mask does not produce additional crowding, especially for positive SOAs, which was also noted for Experiment 3. If there was sufficient masking, crowding with the mask might always be stronger than crowding without the mask. The main finding of the simulations is that the model predicts modest release from crowding (a shine-through effect) that is not supported by the data.

Importantly, the model does exhibit some behavior that is consistent with our broad explanation of crowding for Experiment 4. For example, the model does produce perceptual grouping among the left and right sets of flanking elements, as indicated by illusory contours in Figure 5c at the top and bottom of each side of the flanker gratings, but these contours do not include the mask or target vernier. As argued earlier, such grouping is necessary for release from crowding, but it is not sufficient, as the current model does not have a means for explicitly separating the signals that represent the target from the perceptually grouped signals that represent the flankers; instead, both signals contribute to the template-matching process. For the shine-through effect simulated in Experiment 1, such separation is due to false binocular disparity matches generating target signals in a different depth plane. For the model to account for the findings in Experiment 4, it appears that some new mechanism must be introduced to separate different perceptual groups.

Figure 6a shows that the model does a good job of matching the basic characteristics of the empirical data in Experiment 5. The model generally demonstrates better target discrimination when the mask is present, because the mask allows for a shine-through effect where a representation of the target appears isolated at a near depth plane. This shine-through effect corresponds to a release from crowding. Such a shine-through effect only occurs when there are enough flankers to support perceptual grouping. Figure 6b shows that four flankers are insufficient, and here the model predicts a modest anti-shine-through effect, which partly explains the rise of the curve in Figure 6a. Figure 6c shows that eight flankers do create the illusory contours that correspond to perceptual grouping. Such grouping then enables the shine-through effect and so target discrimination improves. When the mask is absent, the flankers still group, but they do not introduce a shine-through effect, because the lack of a mask precludes the creation of false disparity matches.

Overall, given that the model used the same equations and parameters as Francis (2009), the partial

match between the simulations and empirical data seems promising—especially since we know of no alternative quantitative model that can account for these findings. In particular, Experiments 1, 2, and 5 show threshold patterns that support the model's proposal that release from crowding, in those situations, is the result of a shine-through effect that is engendered by a variety of factors, including perceptual grouping. The model does not perform as well for Experiment 3, although the observed crowding (rather than a release) is generally consistent with the model. The model's worst performance is for Experiment 4, where the model predicts a shine-through effect (release from crowding) but the data do not show it. For reasons that might be accounted for with changed parameters, the model also does poorly for other aspects of Experiments 3 and 4. Overall, the model has some success, and the discrepancies suggest changes that might improve the model's behavior.

To summarize the model limitations, some discrepancies may be due to nonoptimized parameters. For example, the model has mechanisms for lateral inhibition that, in principle, could explain the backward masking effects demonstrated in Experiment 3. Although we cannot guarantee that parameter adjustment would allow the model to fit the data, it seems possible. In contrast, other discrepancies are more fundamental. For example, the weak masking generated by the long flankers in Experiment 4 is entirely contrary to the model's mechanisms (and allowing for backward masking would only further increase the discrepancy). We speculate that the model needs mechanisms that promote separate representations of the grouped flanking elements and the target vernier. We suspect that grouping mechanisms allow the flanking elements to be segmented out from the representation of the target vernier, thereby leaving the vernier essentially isolated. Such isolation would mean that the flanking elements would have little crowding influence on the target. We are exploring whether the model can be extended along these lines.

General discussion

Crowding occurs when a target is flanked by nearby elements. According to Bouma's law, the window of interaction where flankers impair target identification is restricted to half the eccentricity of target presentation (Bouma, 1970). Flankers outside this window have no effect. Because of Bouma's law, crowding research in the last 40 years has mainly focused on single flankers displayed within Bouma's region. Within Bouma's window, crowding is classically explained by pooling and substitution models, which predict more crowding

when the size or number of flankers increases (Balas et al., 2009; Dakin et al., 2010; Freeman & Simoncelli, 2011; Greenwood et al., 2009, 2010; Hanus & Vul, 2013; Huckauf & Heller, 2002; Parkes et al., 2001; Pelli et al., 2004; Pöder, 2012; Strasburger, 2005; Strasburger et al., 1991; van den Berg et al., 2010; Wilkinson et al., 1997). Obviously, these models fail to predict crowding strength with more complex flanker configurations. It seems that, instead, the best predictor of crowding strength is grouping between the target and flankers (Banks, Larson, & Prinzmetal, 1979; Livne & Sagi, 2007, 2010; Malania et al., 2007; Manassi et al., 2012, 2013; Saarela et al., 2009; Saarela et al., 2010; Sayim et al., 2008, 2010; Wolford & Chambers, 1983). Previously, we showed that grouping and *ungrouping* between target and flankers is determined by various factors, namely, similarity (Manassi et al., 2012; Saarela et al., 2009), good gestalt (Manassi et al., 2012), spacing regularity (Manassi et al., 2012; Saarela et al., 2010), and figural aspects (Manassi et al., 2013).

In foveal vision, we found also that pattern completion, another grouping cue, determines crowding (Hermens et al., 2009). Here we first showed that our foveal results also hold true in peripheral vision. As Hermens et al. (2009) did, we combined masking and crowding. Masking did not always lead to a further deterioration of performance, but to *uncrowding*. We suggest that grouping is crucial. When the vernier was flanked by 16 flankers, vernier and flankers grouped by length similarity (Figure 1, black line) or regularity (Figure 3), leading to strong crowding. When the mask of the same length as the vernier was presented simultaneously, crowding strongly decreased compared to the basic crowding condition because the mask groups with the flankers, creating a coherent pattern of similar (Figure 1a) and regular (Figure 3) elements; and thus, the vernier ungroups from the flankers. Hence, when the “picture” is completed, the vernier is released from crowding.

Contrary to models based on *local* interactions between target and flankers, our results show that the completion effect occurs on a *global* level, well beyond Bouma’s window. First, we showed in Experiment 5 that a minimum number of flankers is needed (more than eight) for pattern completion. Second, completion occurs in a much larger window than predicted by Bouma’s law (Bouma, 1970). The vernier target was presented at 3.88° of eccentricity; hence, Bouma’s window is $3.88^\circ/2 = 1.94^\circ$. However, completion occurred only when the sixth and eighth outer flankers were added at 2.33° and 3.11° from the target, respectively. Third, the completion effect occurred only when the flankers formed a regular pattern (Figure 3), i.e., the mask grouped with the *entire* pattern of flankers.

Even though our results in peripheral vision are qualitatively similar to our results with foveally presented stimuli (Hermens et al., 2009), there are some differences. When the pattern of same-length flankers was completed by the mask (Figure 1a), crowding in foveal vision completely disappeared, whereas crowding strength was only halved in peripheral vision. In the fovea, the maximum peak of interaction occurred for an SOA of 0, whereas the lowest and highest thresholds were found for an SOA of 60 ms. It should be noted, however, that Hermens et al. used a pattern of 24 flankers compared to our present study with 16 flankers, so it is difficult to directly compare the two studies.

Here, we combined crowding and visual masking. Traditionally, there are two main types of masking. In A-type masking, performance is worst when target and mask are presented simultaneously (SOA = 0 ms). In B-type masking, performance is worst for SOAs at around 50 ms. In our study, we found evidence for “inverted” A-type masking. Performance was best when the mask and the configuration of flankers and vernier were presented simultaneously (release from crowding). For negative and positive SOAs, performance deteriorated, reaching performance of the unmasked, crowded conditions at SOAs of -200 and $+200$ ms. Clearly, our results cannot be directly compared with most classical masking situations, because we presented a complex target–flanker configuration rather than a single target and mask, as is usual in masking. In addition, we measured rather an indirect measure of target processing, namely, its release from crowding, where the flankers may be seen as a metacontrast mask presented at an SOA of 0 ms. In this sense, our study combines various types of masks, a situation which is usually not addressed by models of masking. Our results are in agreement with other temporal *unmasking* effects: When a target is followed by a first mask which is preceded or followed by a second mask, the two masks can interact with each other, leading to the unmasking of the target (Breitmeyer, Rudd, & Dunn, 1981; Briscoe, Dember, & Warm, 1983; Dember & Purcell, 1967; Ogmen, Breitmeyer, Todd, & Mardon, 2006; Piéron, 1953; Robinson, 1966; Tenkink, 1983; Tenkink & Werner, 1981). Adding elements improves performance in many other paradigms. For example, Pomerantz, Sager, and Stoeber (1977) showed that searching for a target is faster when additional elements lead to an emergent feature and, thus, the target pops out. Along the same lines, Pomerantz and Schweitberg (1975) showed that grouping between two stimuli can be eliminated by introducing a third element.

Our results showed that the combination of crowding and masking can lead to a decrease of crowding. For a large pattern of elements with equal-

length mask and flankers (Figure 1a), the vernier was perceived as brighter for short SOAs (0, 20, and 60 ms) compared to the conditions with longer SOAs (200 ms). This effect is reminiscent of the shine-through masking effect (Herzog, Dependahl, Schmonsees, & Fahle, 2004; Herzog & Koch, 2001). The results from Experiment 5 provide further evidence in this direction: The pattern completion effect did not occur with two, four, or eight flankers, but only with a larger pattern of elements. Vickery, Shim, Chakravarthi, Jiang, and Luedeman (2009) found that the combination of crowding and masking can also lead to a decrease of performance. Orientation discrimination of a letter T only slightly deteriorated when either flanking Ts were presented (crowding condition) or a square surrounded the target (masking condition). When the two conditions were combined, super-crowding occurred.

We argued that existing models of crowding could not explain the results reported here because they lack a mechanism for perceptual grouping. We then explored a neural-network model that contains mechanisms to deal with perceptual grouping and that was previously able to account for a variety of effects related to backward masking (Francis, 1997) and the shine-through effect (Francis, 2009). We found promising but mixed outcomes. In the model, crowding can be influenced by perceptual grouping, but only when such grouping generates a shine-through effect that leads to a neural representation that isolates the vernier from its surround. Although this property matches the findings in Experiments 1, 2, and 5, it does not agree with the findings in Experiments 3 and 4. The model's behavior for Experiment 4 wrongly predicted that perceptual grouping of the flankers without a shine-through effect leads to strong crowding. This discrepancy indicates that the model needs additional mechanisms that segment visual representations as a result of perceptual grouping.

In conclusion, our results show once again that crowding models, like simple pooling and substitution, have limitations in predicting crowding strength (see Herzog & Manassi, 2015, for in-depth discussion). Our data suggest that perceptual grouping determines crowding strength. Although grouping does not explain why target perception deteriorates in crowding, our results show that perceptual organization plays a crucial role in predicting crowding strength. Perceptual organization has been shown to play a role also in many other visual processes, like surround suppression (Saarela & Herzog, 2009a), metacontrast masking (Duangudom, Francis, & Herzog, 2007), visual short-term memory (Kahneman, 1973), and audition (Bregman, 1981; Oberfeld, Stahn, & Kuta, 2014).

Keywords: crowding, masking, grouping, contextual modulation, pattern, vernier acuity, completion, modeling

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. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 604102 (HBP). Frouke Hermens is now at the University of Lincoln (UK).

Commercial relationships: none.

Corresponding author: Mauro Manassi.

Email: mauro.manassi@berkeley.edu.

Address: Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

References

- Andriessen, J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research*, 16(1), 71–78, <http://www.sciencedirect.com/science/article/pii/004269897690078X>, doi:10.1016/0042-6989(76)90078-X.
- Bach, M. (1996). The Freiburg Visual Acuity Test—Automatic measurement of visual acuity. *Optometry and Vision Science*, 73(1), 49–53.
- 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, doi:10.1167/9.12.13. [PubMed] [Article]
- Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics*, 25(6), 447–456.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226(5241), 177–178, <http://www.ncbi.nlm.nih.gov/pubmed/5437004>.
- Bregman, A. S., (1981). Asking the “what for” question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Lawrence Erlbaum.
- Breitmeyer, B. G., Rudd, M., & Dunn, K. (1981). Metacontrast investigations of sustained-transient

- channel inhibitory interactions. *Journal of Experimental Psychology, Human Perception and Performance*, 7(4), 770–779.
- Briscoe, G., Dember, W. N., & Warm, J. S. (1983). Target recovery in visual backward masking: No clear explanation in sight. *Journal of Experimental Psychology, Human Perception and Performance*, 9(6), 898–911.
- Cao, Y., & Grossberg, S. (2005). A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis. *Spatial Vision*, 18(5), 515–578, <http://booksandjournals.brillonline.com/content/journals/10.1163/156856805774406756>.
- Dakin, S. C., Cass, J., Greenwood, J. A., & Bex, P. J. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision*, 10(10):14, 1–16, doi:10.1167/10.10.14. [PubMed] [Article]
- Dember, W. N., & Purcell, D. G. (1967). Recovery of masked visual targets by inhibition of the masking stimulus. *Science*, 157(3794), 1335–1336.
- 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, doi:10.1167/7.1.7. [PubMed] [Article]
- Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, 104(3), 572–594.
- Francis, G. (2009). Cortical dynamics of figure-ground segmentation: Shine-through. *Vision Research*, 49(1), 140–163, doi:10.1016/j.visres.2008.10.002.
- Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, 14(9), 1195–1201.
- 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, USA*, 106(31), 13130–13135, <http://www.pnas.org/content/106/31/13130.abstract>, doi:10.1073/pnas.0901352106.
- 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>.
- Hanus, D., & Vul, E. (2013). Quantifying error distributions in crowding. *Journal of Vision*, 13(4):17, 1–27, doi:10.1167/13.4.17. [PubMed] [Article]
- Hermens, F., Herzog, M. H., & Francis, G. (2009). Combining simultaneous with temporal masking. *Journal of Experimental Psychology, Human Perception and Performance*, 35(4), 977–988.
- Herzog, M. H., Dependahl, S., Schmonsees, U., & Fahle, M. (2004). Valences in contextual vision. *Vision Research*, 44(27), 3131–3143.
- Herzog, M. H., & Fahle, M. (2002). Effects of grouping in contextual modulation. *Nature*, 415(6870), 433–436.
- Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proceedings of the National Academy of Sciences, USA*, 98(7), 4271–4275.
- Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: A fresh look at object recognition. *Current Opinion in Behavioral Sciences*, 1(0), 86–93, doi:10.1016/j.cobeha.2014.10.006.
- Huckauf, A., & Heller, D. (2002). Spatial selection in peripheral letter recognition: In search of boundary conditions. *Acta Psychologica*, 111(1), 101–123.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, 48(5), 635–654, doi:10.1016/j.visres.2007.12.009.
- Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision*, 7(2):4, 1–12, 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, doi:10.1167/10.3.1. [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, 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, 1–14, doi:10.1167/12.10.13. [PubMed] [Article]
- Manassi, M., Sayim, B., & Herzog, M. H. (2013). When crowding of crowding leads to uncrowding. *Journal of Vision*, 13(13):10, 1–10, doi:10.1167/13.13.10. [PubMed] [Article]
- Oberfeld, D., Stahn, P., & Kuta, M. (2014). Why do forward maskers affect auditory intensity discrimination? Evidence from “Molecular Psychophysics.” *PLoS One*, 9(6), e99745, doi:10.1371/journal.pone.0099745.
- Ogmen, H., Breitmeyer, B. G., Todd, S., & Mardon, L. (2006). Target recovery in metacontrast: The effect of contrast. *Vision Research*, 46(28), 4726–4734, doi:10.1016/j.visres.2006.09.006.
- 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, 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>
- Piéron, H. (1953). Le processus du metacontraste. *Journal de Psychologie*, 32, 1–24.
- Pöder, E. (2012). On the rules of integration of crowded orientation signals. *i-Perception*, 3(7), 440–454.
- Pomerantz, J. R., Sager, L. C., & Stoeber, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology, Human Perception and Performance*, 3(3), 422–435.
- Pomerantz, J. R., & Schweitzerberg, S. (1975). Grouping by proximity: Selective attention measures. *Perception & Psychophysics*, 18(5), 355–361.
- Robinson, D. N. (1966). Disinhibition of visually masked stimuli. *Science*, 154(3745), 157–158.
- Saarela, T. P., & Herzog, M. H. (2009a). Size tuning and contextual modulation of backward contrast masking. *Journal of Vision*, 9(11):21, 1–12, 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, 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, doi:10.1167/10.10.17. [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, 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, 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.
- Tenink, E. (1983). Recovery at short intervals between masking flashes. *Vision Research*, 23(12), 1693–1698.
- Tenink, E., & Werner, J. H. (1981). The intervals at which homogeneous flashes recover masked targets. *Perception & Psychophysics*, 30(2), 129–132.
- van den Berg, R., Roerdink, J. B. T., & Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology*, 6, e1000646.
- 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, doi:10.1167/9.2.12. [PubMed] [Article]
- Wichmann, F. A., & Hill, N. J. (2001. Nov). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8), 1293–1313.
- 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. Feb). Lateral masking as a function of spacing. *Perception & Psychophysics*, 33(2), 129–138.

Appendix: Modeling details

Details of the model equations and parameters are given by Francis (2009). The only changes to the simulations were to generate stimuli that matched those used in Experiments 1–5 (the model simulations include only horizontal and vertical orientations, so they cannot consider the stimuli used in Experiment 6, which include off-axis orientations). For the simulations, we oversampled the SOA values so that we could be sure of properly characterizing the model's behavior. Model performance was derived from template matches for a vernier shifted to the left, M_L , or right, M_R . The templates were always centered on the target vernier and uniformly summed signals within 30×40

pixels for a region covering the top left and another region covering the bottom right of the target. The width is 5 times the spacing between stimulus elements, and the height is twice the length of the target vernier offset line. Such a large template allows for crowding when flanking elements influence template matching. Energy for the vernier direction being a shift to the right at time t was then computed as a contrast

$$C_R(t) = \frac{M_R(t) - M_L(t)}{0.01 + M_R(t) + M_L(t)}, \quad (1)$$

where the constant 0.01 avoids division by 0. As described by Francis (2009), the model includes two distinct depth planes, near and far. The template matches are computed for both depth planes, with contrast energies indicated by $C_{Rn}(t)$ and $C_{Rf}(t)$ for the near and far planes, respectively. The evidence for the

vernier being shifted to the right was an integration of contrast energy across time for both depth planes:

$$E_R = \int_{\tau_1}^{\tau_2} C_{Rn}(t) + C_{Rf}(t) dt, \quad (2)$$

where τ_1 is the start of the trial and τ_2 is the end of the simulation for a trial (identified as when the target signals disappear). Larger values of E_R indicate better detection of the vernier direction and thus smaller thresholds. For all model simulation plots, the y-axis is plotted in reverse for easy comparison with the experimental data. C++ source code for the model is available at https://osf.io/ey9jz/?view_only=9bcee713a79a4dc09aa8393637cee73b. By default, the simulation reproduces model results for all of the plots reported in this paper.