The same binding in contour integration and crowding

Ramakrishna Chakravarthi and
Université de Toulouse, CerCo, UPS, France, CNRS, UMR 5549, Faculté de Médecine de, Rangueil, Toulouse France

Denis G. Pelli
Psychology and Neural Science, New York, University, New York, NY, USA

Abstract

Binding of features helps object recognition in contour integration, but hinders it in crowding. In contour integration, aligned adjacent objects group together to form a path. In crowding, flanking objects make the target unidentifiable. But, to date, the two tasks have only been studied separately. May and Hess (2007) suggested that the same binding mediates both tasks. To test this idea, we ask observers to perform two different tasks with the same stimulus. We present oriented grating patches that form a "snake letter" in the periphery. Observers report either the identity of the whole letter (contour integration task) or the phase of one of the grating patches (crowding task). We manipulate the strength of binding between gratings by varying the alignment between them, i.e. the Gestalt goodness of continuation, measured as "wiggle". We find that better alignment strengthens binding, which improves contour integration and worsens crowding. Observers show equal sensitivity to alignment in these two very different tasks, suggesting that the same binding mechanism underlies both phenomena. It has been claimed that grouping among flankers reduces their crowding of the target. Instead, we find that these published cases of weak crowding are due to weak binding resulting from target-flanker misalignment. We conclude that crowding is mediated solely by the grouping of flankers with the target and is independent of grouping among flankers.

Keywords

crowding; wiggle; grouping; binding; Gestalt; contour integration; good continuation; alignment; object recognition; snake letter

Introduction

The Gestalt law of good continuation, the tendency to perceive aligned adjacent objects as belonging to a group, has been extensively studied (Wertheimer, 1923; Geisler, 2008; Loffler, 2008; Smits, Vos, & van Oeffelen, 1985; Field, Hayes, and Hess, 1993; Kovacs and Julesz, 1993; Hess and Field, 1999; Hess and Dakin, 1999). One example of this binding of objects into a group is contour integration, which is usually assessed by tasks that involve detecting the presence of a chain of objects forming a contour in a field of randomly oriented objects: a "snake in the grass". Binding the aligned objects helps in that task, but binding can also hinder. In peripheral vision, an otherwise identifiable object becomes unidentifiable in the presence of nearby flankers. This phenomenon of interference between

target and flankers is known as crowding (Stuart and Burian, 1962; Bouma, 1970). It seems that the flanker and target are bound together to form a jumbled unidentifiable whole. Thus, we have two phenomena in which flankers play opposite roles. In contour integration, bound flankers help, whereas, in crowding, bound flankers hinder. In this study, we investigate whether the same binding process links objects in these two disparate phenomena, as suggested by May & Hess (2007). We also reexamine the suggestion by Livne and Sagi (2007) that binding among distracters affects the strength of their binding to the target.

Until recently, all studies of contour integration have focused on the detection of a contour, reporting its presence or absence (or 2afc), and not on identification, choosing one among many possible identities. Unlike detection, identification typically involves quickly categorizing a stimulus into one of many possibilities that are familiar, meaningful, and nameable. A recent study goes beyond detection to establish the first direct link between good continuation and object recognition. The efficiency of identifying letters made up of gabors increased with the goodness of continuation between the gabors: efficiency was inversely proportional to "wiggle," a measure of misalignment of the gabors (Pelli, Majaj, Raizman, Christian, Kim, & Palomares, 2009). That is, the better the gabors line up, the easier it is to identify the letter. Here we extend that finding by exploring whether alignment also plays a role in crowding, a breakdown of object recognition.

Models of object binding attempt to account for the results of contour integration tasks. According to the popular "Association Field" model, the strength of binding between adjacent objects depends on their separation and alignment (Field et al. 1993). It has been argued that contour detection performance can be explained by such local binding processes (Field et al. 1993; but see Loffler, 2008).

In crowding, nearby flankers reduce the identifiability of a target in the peripheral visual field (Bouma 1970). This effect depends on the separation between the flankers and the target. *Critical spacing* is the center-to-center distance between the target and the flankers beyond which the flankers do not affect target identification. Critical spacing is proportional to eccentricity (Bouma, 1970; Toet and Levi, 1992; Pelli and Tillman, 2008). Most theories of crowding hold that the difficulty in identification arises from excess binding that inappropriately combines flanker features with those of the target (Pelli, Palomares, and Majaj, 2004; Levi, 2008; Parkes, Lund, Angelucci, Solomon, and Morgan, 2001; Levi, Hariharan, and Klein, 2002). That is, the features of the target and the distracters are bound together if the separation between them is less than the observer's critical spacing for that eccentricity. We noted that alignment matters in contour integration. Does it matter in crowding as well?

The Gestalt law of similarity, the tendency to perceive similar objects as belonging to a group, is important in crowding (Kooi, Toet, Tripathy, and Levi, 1994). Thus, two Gestalt laws – good continuation and similarity – postulate binding between the target and flankers that would occur in crowded situations. Figure 1 demonstrates both: alignment and similarity. Here, we investigate how the Gestalt law of good continuation affects binding in two tasks.

Some argue that the two phenomena – crowding and contour integration – could be related (May and Hess, 2007, Livne and Sagi, 2007, Dakin, Cass, Greenwood, and Bex, 2010). May and Hess (2007) suggested that the association field in contour integration and the combining field in crowding (the area within critical spacing) might be one and the same. Developing that idea, they showed that a crowding-based model could explain why, in the periphery, it is so much harder to see ladder contours than snake contours. Here, we directly test that proposal.

Flankers play opposite roles in contour integration and crowding. Here we examine the two phenomena through two tasks performed with the same stimulus, applying the same stimulus manipulation to both tasks. Finding that the threshold for binding in both tasks is affected in the same way by this manipulation would indicate that the same or similar binding mechanisms underlie both phenomena.

Experiment

To test whether the same binding process links objects together in contour integration and in crowding, our experiment uses the same stimulus to probe both phenomena. We perturb the alignment between gabors in order to assess the effect of the strength of binding on contour integration and crowding. If the same binding process underlies both, strengthening the link between gabors will improve contour integration and impair identification of a particular gabor (i.e., increase crowding). We arrange gabors to make a letter (Fig. 2). The center of this letter is 10 deg to the right of fixation. Observers perform one of two tasks with this stimulus. The *contour integration task* asks them to identify the letter. The *crowding task* asks them to identify the phase of the central gabor (i.e., to indicate if the light half of the gabor was on the right side or the left). The relative orientation between gabors is changed to vary the goodness of continuation (alignment) between them. One advantage of using a letter identification task instead of the usual "snake in the grass" detection is that, by testing identification rather than detection, we can draw conclusions about object recognition. The contour integration task employed here is similar to the Pelli et al. (2009) "snake letter" identification task, which measured the effect of alignment on object recognition.

Identifying snake letters seems to be a good test of contour integration. But might it be contaminated by crowding among the elements, as occurs among the parts of a face (Martelli, Majaj, and Pelli, 2005)? No. First, specifically, faces (and words) have parts but letters don't (Martelli et al. 2005), so local elements within a letter do not crowd the identity of the letter itself. Second, more generally, Pelli et al. (2009) found similar effects of wiggle on letter recognition in the fovea (i.e. without crowding) as found here in the periphery. Thus, the known facts continue to endorse identification of snake letters as an assay of contour integration.

Methods

Observers

Three experienced observers (including the first author), aged 25 – 32, with normal or corrected-to-normal vision took part in this experiment.

Stimuli

Stimuli are generated using MATLAB with the Psych-toolbox extensions (Brainard, 1997; Pelli, 1997) running on an Apple G4 Macintosh computer and presented on an 18 inch CRT monitor with a resolution of 1024 x 768 pixels and a frame rate of 100 Hz. The display is placed 57 cm from the observer, whose head is stabilized with a chin and forehead rest. At this distance there are 29 pixels/deg.

The stimulus (Fig. 2) consists of gabors presented on a uniform gray background with luminance 42 cd/m^2. Each gabor is a sinewave grating vignetted by a Gaussian envelope. The contrast function for a vertical gabor at fixation is

$$g(x,y)=\sin(\psi+2\pi f x)\exp\frac{-\left(x^2+y^2\right)}{\lambda^2}, \quad (1)$$

where ψ is phase offset, $f = 1$ c/deg is spatial frequency, x and y are horizontal and vertical position in deg, and $\lambda = 0.37$ deg is the space constant of the envelope. We will later refer to the *dark-to-light transition* of this function as the line in *x-y* space at which the argument of sin( ) is zero. (When we fit a sine to this, to measure "wiggle", the contact is at the point along the transition line where the envelope exp( ) is maximum.) The luminance function for vertical gabors at *n* locations is

$$L(x,y)=\left(1+\sum_{i=1}^{n}g(x-x_i,y-y_i)\right)42\,\text{cd/m}^2, \quad (2)$$

where x_i, y_i is the position of the *i*-th gabor. The gabors are presented in a grid, 3 horizontally and 5 vertically, to form a letter. The center of the grid is 10 deg to the right of fixation. The center-to-center separation between adjacent gabors is 1.5 deg.

On each trial we build the stimulus as follows. First, for the given letter being presented in that trial, we place gabors along the path of the letter. The orientations of the gabors are all tilted to the same extent ($\theta = 0, 20, 40, 60,$ or 90 deg) with respect to the letter path. However, the direction of the tilt alternates between adjacent gabors ($\pm\theta$ with respect to the letter path). An extra orientation jitter, randomly chosen from a uniform distribution (-5 to 5 deg), is applied to each gabor on each trial. The variation in relative orientation of adjacent gabors controls the goodness of continuation (alignment) between them. Following Pelli et al. (2009), we quantified the misalignment as *wiggle*. The smaller the wiggle, the better the alignment. To calculate a letter's wiggle, we identify straight chains of gabors in the letter. Within each chain, for each pair of gabors adjacent to each other, we fit a one-half sine wave to make tangential contact (at each end) with the dark-to-light transition of each gabor (Fig. 3). At the point of contact, the dark side of the transition is always to the right of the sine as the sin( ) argument increases.

The period of the sine wave is set to twice the center-to-center spacing of the two adjacent gabors. The sine ends at the contact points described above (the point along the transition line where the Gaussian envelope is maximum). The position, orientation, amplitude and phase of the wave are adjusted so as to produce the best fit (i.e., minimizing the angle between the tangent to the sine wave at the contact point and the transition line). The angle that this sine wave makes with its own axis is that pair's *wiggle*. We repeat this for each pair of adjacent gabors. We then average the wiggles of all gabor pairs in all the straight chains in the letter, to obtain the letter's *wiggle* (Fig. 3). This procedure is a generalization of the Pelli et al. (2009) method of measuring the wiggle of periodic stimuli to include some non-periodic stimuli.

Procedure

There are two tasks: *crowding* and *contour integration*. Each observer participated in both. The stimulus parameters and procedure are the same in both, except where noted. Each task has five conditions, which differ only in the orientation θ of the component gabors relative to the letter path. Wiggle increases with θ. We ran four blocks of 40 trials per condition. The order of blocks (task type by wiggle angle) is randomized for each observer. Each block begins with a press of the space bar. A black fixation square appears at the center of the screen throughout the whole block. Viewing is binocular. The many-gabor stimulus is

presented in the right field for 150 ms. The stimulus is the capital letter I in the crowding task and one of the capital letters C, E, F, H, I, L, O, P, T, or U in the contour integration task. (This difference in letters—"I" versus others—between tasks is inconsequential, as shown in the Results section.) The observer then indicates his or her response with a key press. The crowding task is to report the phase of the target gabor – the gabor in the center of the grid (i.e., the center of the letter I) — in other words, to indicate the left/right location of the light half of the gabor. The phase of this target gabor is random ($\psi = 0°$ or $180°$) on each trial. The rest of the gabors have zero phase. The observer indicates the target's phase, after the stimulus is taken away, by pressing either the left (or up) or the right (or down) arrow key. The left and up arrow keys are equivalent, and the right and down arrow keys are equivalent. The contour integration task is to report the identity of the letter, which is selected independently, randomly, for each trial. In this task all the gabors have zero phase. A correct response is rewarded with a low-frequency beep, whereas an incorrect response is indicated by a high-frequency beep. The next trial is presented after an intertrial interval of 1 s.

Performance in each task is tested at orientations θ of 0, 20, 40, 60, and 90 deg. In the crowding task, the phase of the target gabor is random on each trial: either 0 or 180 deg. That is, the target is phase-aligned with the flankers in the same-target-phase trials (target phase 0 deg) and misaligned in the opposite-target-phase trials (target phase 180 deg). The phase difference results in different amounts of wiggle in these two types of trials, so we analyze the same- and opposite-target-phase trials separately and plot the results at their respective wiggles.

The QUEST algorithm (Watson and Pelli, 1983) is used to determine the contrast threshold (82% correct) for each of the tasks. Our threshold estimate (log contrast) is the mean of the posterior probability density function based on a Weibull fit to the data. To avoid the possibility of surround suppression in the crowding task (and to maintain similar stimulus characteristics between the two tasks) we use equal contrast for all gabors (Petrov, Popple, and McKee, 2007). We also measure the contrast threshold for phase discrimination of a single gabor, in the absence of flankers, to provide a baseline for the crowding task.

For each observer, four contrast thresholds (one from each block) were obtained for each orientation in each task. A block was discarded if the standard deviation of the threshold as estimated by QUEST was too large (> 0.5 log units of contrast) or if the estimated threshold was outside the range of tested contrast values. We ran a set of 4 additional blocks with stimuli of the same orientation in such cases, retaining all threshold estimates with acceptable standard deviations. Thus, observer SBR participated in a total of 44 blocks of which 2 blocks were discarded; LH participated in 44 blocks of which 2 blocks were discarded; CRK participated in 60 blocks of which 7 blocks were discarded. We report the average log contrast thresholds for each condition (task and wiggle).

In the crowding task, for a given orientation θ, there are two kinds of trials: trials with target phase 0 deg and trials with target phase 180 deg. These two types of trials have different wiggles (Fig. 3). Since we are interested in the effect of wiggle on crowding, we analyze data from the trials with different target phases separately. The crowding task was to identify the phase of the target, so the 0 and 180 deg target-phase trials must be interleaved. Therefore, to obtain contrast thresholds for each phase we adopted the following procedure. We collected all the trials from all blocks at each θ. We then separated these trials into two groups: 0 and 180 deg phase.. Each group had roughly 80 trials. For each group we obtained two contrast thresholds by randomly picking half the trials and using QUEST to obtain a threshold for each half. Thus for the crowding task, we report the log average of two contrast thresholds per wiggle.

To reiterate the hypothesis, if the same binding mechanism underlies both contour integration and crowding, then we would expect similar but opposite effects of wiggle on the contrast thresholds in the contour integration and crowding tasks.

Results

Results for both tasks are plotted in Figure 4 for each of the three observers. The effect of wiggle on contrast threshold for each task is monotonic for wiggle values up to 60 deg. We discuss this monotonic effect first and then consider the nonmonotonic result at 90 deg wiggle. As can be seen in these plots, increasing wiggle (reducing alignment) up to 60 deg makes it harder to identify a letter composed of gabors and easier to identify the phase of a single target gabor among flankers. Increasing wiggle impairs contour integration and relieves crowding. These results suggest that the same or similar binding processes underlie both crowding and grouping by alignment. The increase in threshold for letter identification with increasing wiggle replicates the existing central-field results (Pelli et al. 2009), but in the peripheral visual field. As mentioned in the Methods section above, the wiggles for the opposite-target-phase trials in the crowding task are larger (to the right along the horizontal axis) than those for the same target phase. We plot each contrast threshold at its wiggle. The thresholds show the same dependence on wiggle in both kinds of trials.

The measured effect of wiggle on the two tasks tests a particularly simple version of our hypothesis. Suppose the two tasks have the same binding mechanism and that binding of the flankers is all or none. That is, each trial is perceptually bound (or unbound), with a probability B that depends solely on the amount of wiggle. We suppose that at zero wiggle all trials are bound, $B=1$, and that at 60 deg wiggle all trials are unbound, $B=0$. We call the amount of wiggle that produces 50% probability of binding ($B=0.5$) the wiggle threshold W_{50}. If both tasks involve the same binding then they must have the same wiggle threshold. The following six steps explain how we arrive at and test this prediction. First, the expected proportion correct for each task depends on contrast c and on whether the trial is bound or unbound. Lacking complete models for how the two tasks are performed, we cannot say how much difference the binding should make, and the dependence of the probability of binding on wiggle may be nonlinear and even non-monotonic. However, the expected fraction of trials that are bound is the probability of binding B, so the measured proportion correct P (including a mixture of bound and unbound trials) will be a linear interpolation (0 to 100%) between the proportions correct at zero and 60 deg wiggle, $P = (P_0{}^\circ - P_{60}{}^\circ) B + P_{60}{}^\circ$. Thus proportion correct P is linearly related to probability of binding B. Second, the P vs. log c psychometric function, as a whole, is nonlinear (sigmoidal), but it is smooth, so we suppose that, over a modest range, proportion correct is linearly related to log contrast. Third, the cascade of several linear relations is itself a linear relation, so, provided P is near some criterion value (e.g. 0.82), log contrast threshold is linearly related to probability of binding. Fourth, the results in Fig. 4 are all collected at a fixed threshold criterion P, satisfying the third step's proviso. Fifth, although the linear relation is task-dependent, the linearity itself implies that the midway point of binding probability, $B=0.5$, corresponds to the midway point of log contrast threshold, log $c = (\log c_0{}^\circ + \log c_{60}{}^\circ)/2$. Thus, sixth, if the same binding process mediates both tasks, we predict that the wiggle threshold W_{50} will be the same for both tasks. Figure 5 is a scatter plot of each observer's wiggle threshold for the two tasks: contour integration vs. crowding. As predicted, all the data points lie close to equality (diagonal line). The conjectured linearity implies that the wiggle thresholds for each task should be around 30 deg (the midpoint between 0 and 60 deg), which is what we find.

We summarize the effect of wiggle by calculating the wiggle threshold W_{50} for each task and observer (Fig. 5). The dashed curves are quadratic polynomials fitted to each observer's results (log contrast threshold vs. wiggle, 0 to 60 deg). A single polynomial is fitted to same-

and opposite-target-phase thresholds in the phase discrimination task. To estimate the observer's wiggle threshold W_{50}, we find the amount of wiggle at which that curve's log contrast is midway between the log contrast thresholds at wiggles of 0 and 60 deg. The predicted equality of wiggle thresholds across tasks rests on linearity. As explained in the second paragraph of Results, all tasks for which log c is linearly related to wiggle will have the same wiggle threshold: 30 deg. Thus it is an essential feature of our test that we allow for a nonlinear relation by making a quadratic polynomial fit.

We did not anticipate the non-monotonicity of the dependence of contrast threshold on wiggle, but one of our anonymous reviewers did. Beyond 60 deg wiggle, the curves turn back toward the zero-wiggle baseline (Fig. 4). Originally, we had measured only the monotonic range, up to 60 deg wiggle. The reviewer suggested collecting data at 90 deg, predicting less binding. This prediction is supported by findings in several studies (Field et al., 1993; Ledgeway, Hess, and Geisler, 2005; May and Hess, 2007). We find that, indeed, letter identification is easier (and phase discrimination is harder) at 90 deg wiggle than at 60 deg in 2 of 3 observers. The variation among observers (e.g. one observer shows no change in performance when wiggle is increased from 60 to 90 deg) is predicted by May and Hess's (2007) model, where the strength of binding in ladders relative to snakes can vary between observers. The precisely opposite results for each observer for contour integration vs. crowding support the suggestion that the same binding underlies both phenomena.

Statistics

An ANOVA was performed on each observer's results separately with log contrast threshold as the dependent measure and task and wiggle as the (independent) factors. For this analysis we considered data for wiggles between 0 and 60 deg. Further, for the crowding task, the ANOVA was computed using data from the same-target-phase trials, for two reasons: i) the wiggles in both tasks are precisely the same, and ii) the effect of wiggle on thresholds seems to be the same for both target-phase trials. In agreement with the latter point, we found similar results when we performed the same ANOVA using data from opposite-target-phase trials.

The main effect of task was significant in all three observers [SBR: $F(1,16)=89.2$, $p<0.0001$; LH: $F(1,16)=39.4$, $p<0.0001$; CRK: $F(1,19)=96.4$, $p<0.0001$]. That is, it was easier to identify a letter made of gabors than it was to indicate the phase of a single gabor among flankers. Importantly for our hypothesis, the interaction between task and wiggle was highly significant in all three observers [SBR: $F(3,16)=12.8$, $p<0.0005$; LH: $F(3,16)=12.7$, $p<0.0005$; CRK: $F(3,19)=13.2$, $p<0.0005$].

Two potential confounds

Above, we claim to be using the "same" stimulus for both tasks, but, in fact, we use one of many letters (C E F H I L O P T U) in the contour integration task and only one (I) in the crowding task. Furthermore, the task-relevant information is restricted to just the target gabor in one of the tasks and distributed over all the gabors in the other task. Obviously, these two differences would be a concern if the two tasks had turned out to have different wiggle thresholds because the stimulus differences might account for the difference in wiggle thresholds. However, in fact, we found practically the same wiggle threshold in both tasks (Fig. 5). It seems unlikely that the task, alphabet, and distribution of information have large effects that just happen to cancel out to zero. This suggests that none of these differences affected sensitivity to wiggle, but we checked anyway, to make sure.

To check for an effect of letter differences, we analyzed only those trials (pooled across all blocks with the same wiggle) in the letter identification task where the target letter was "I".

We obtained practically the same results with just the letter-I trials (Fig. 4, solid triangles) as with all the trials (solid squares). Thus, the differences between "I" and the rest of the letters did not affect wiggle sensitivity.

To check for an effect of the distribution of task-relevant information on wiggle sensitivity, we measured the effect of extending the region containing task-relevant information in the crowding task. Observers reported the phase of either one or three gabors (presented on an imaginary vertical line). These target gabors were embedded among other gabors to ensure crowding. Would changing the extent of task-related information (from one to three gabors) affect performance? We tested performance, in two observers, at 3 wiggle values between 0 and 90 deg. We found that the effect of wiggle was the same as in the main experiment irrespective of the number of gabors to be reported (Fig. 6). As seen earlier, performance was best at intermediate wiggle values and worst at the lowest wiggle (fully crowded), with performance at 90 deg being either intermediate or equal to that at the lowest wiggles tested. In other words, extending the task-relevant information over a larger region had no effect on the pattern of performance in the crowding task.

Thus we rule out both potential confounds. The differences in letters and distribution of information don't affect wiggle sensitivity. Our two tasks, crowding and contour integration, have equal sensitivity to wiggle.

Oblique effect?

In principle, one might try to explain our results by supposing that both tasks (contour integration and crowding) are affected similarly by a third underlying factor. At zero wiggle, in our letters, all the gabors are either horizontal or vertical (Fig. 2). The gabors become more oblique with increasing wiggle. Oblique gabors are known to be weaker stimuli than gabors in the cardinal (vertical and horizontal) orientations (Appelle, 1972; Westheimer, 2003). Therefore, oblique gabors might integrate less well in the contour integration task, making it harder to identify letters, and might be less effective flankers in the crowding task, making it easier to identify the phase of a target gabor. Might the effect we see result from the weakness of oblique gabors as stimuli? No. First, the oblique effect is too weak to mediate the effect of wiggle. It is equivalent to an increase in contrast thresholds by a factor of square root of 2 (Campbell, Kulikowski, and Levinson, 1966), whereas the effect of wiggle is more than a factor of 2. Second, to test this idea, we rotated the entire snake-letter stimulus 45 deg clockwise from vertical, so that gabors in the stimulus at zero wiggle were obliquely oriented. We asked observers to perform the same two tasks as before and found that the thresholds were little affected by this manipulation (see Fig. 4). This rules out the oblique effect as an explanation for the measured effect of wiggle, though it might make a minor contribution to a very precise account.

Discussion

Alignment (good continuation) improves letter recognition and impairs phase discrimination. Observers have equal sensitivity to wiggle in these two very different tasks. This suggests that the same or similar binding processes underlie contour integration and crowding.

Flankers make a peripheral target unrecognizable. This phenomenon of crowding is a breakdown of object recognition. Most theories of crowding propose that this problem arises at the feature combination stage, after the feature detection stage (Pelli et al. 2004; Levi, 2008; Pelli and Tillman, 2008). The idea is that all features detected within the space (combining field) circumscribed by the critical spacing at a given eccentricity are bound together. If the features belong to different objects, as occurs when flankers are within a

critical spacing of the target, the target is crowded. We find that the contrast threshold for phase discrimination of a gabor embedded among other gabors decreases with increasing wiggle, showing that the binding (assessed by crowding) between target and flanker is stronger if they are aligned. This exposes the role of goodness of continuation in crowding and object recognition.

Binding in crowding

Binding is an essential part of object recognition. Over most of its history, crowding (under various names) has been described as a pernicious process afflicting amblyopic and peripheral vision. It is widely accepted that crowding is unwanted binding, the inappropriate binding of target features with flanker features (Palomares, LaPutt, & Pelli, 1999; Levi et al. 2002; Parkes et al. 2001; Pelli et al. 2004; Intriligator & Cavanagh, 2001; Levi 2008). Here we confirm the suggestion that crowding may be just a label, not necessarily a separate process. The same binding that is called "good" when object recognition succeeds (e.g. seeing the snake letter) is called "crowding" when object recognition fails (e.g. failing to identify the phase of a target gabor).

In particular, Parkes et al. (2001) suggested that features (within the critical spacing) are pooled without regard to location. However, Livne and Sagi (2007) show an example in which rearranging flankers so that they group with each other drastically weakens crowding. That is, simple pooling of the target with flankers, without regard to arrangement, is not the whole story. We agree, but they went on to conclude that the reason that arrangement matters is that flankers that are grouped together are less effective in crowding the target. Here, we will argue, instead, that, in all these cases, flanker-to-flanker grouping is irrelevant and that crowding depends solely only on the binding (i.e. grouping) of target to flankers.

Let us reexamine the Livne and Sagi (2007) results. The configurations that raised threshold for the target (i.e., resulted in crowding) in their experiments were those in which the target grouped with at least some of the flankers. Figure 7 reproduces the stimuli that they used to reach the conclusion that, when flankers are grouped with each other, crowding is reduced (no crowding in A and strong crowding in B). If their conclusion is right, then any break in the contour should increase crowding, because some flankers will be misaligned, weakening flanker-to-flanker grouping. Alternatively, if our interpretation is right, then crowding is stronger if flankers group with the target, regardless of whether the flankers are grouped with each other. We predict that if the target is aligned with any distracters, then it will be strongly crowded, regardless of whether the flankers are aligned with each other. Thus, according to their hypothesis, configurations C and D should be equivalent. However, the target in C is much less well aligned with flankers than in D, so we predict that C should produce no or minimal crowding and D should produce significant crowding, which, in fact, is what they found.

Further evidence is provided by three recent studies. Crowding is stronger when flankers group with the target by virtue of similarity (Saarela, Sayim, Westheimer and Herzog, 2009) or by regularity of spacing (Saarela, Westheimer and Herzog, 2010) than when they are not grouped. Target-flanker grouping overrides any effect of grouping among flankers in determining performance under crowded conditions. Similarly, Yeotikar, Khuu, Asper, and Suttle, (2011) found that crowding is stronger when flanking gabors are aligned with the target than when they are orthogonal to it. Other studies have reported that flankers parallel to the target raise threshold much more than perpendicular flankers do (Andriessen & Bouma, 1976; Wilkinson, Wilson, & Ellemberg, 1997). Flankers group with each other in both cases, but thresholds are raised only when the flankers group with the target.

Binding in contour integration

May and Hess (2007) reported a difference in detecting two kinds of contour in the periphery. In a "grass" field of randomly oriented gabors, it is easy to detect a "snake" contour made of gabors aligned with its path, but hard to detect a "ladder" contour made of gabors oriented perpendicular to its path. We compute the wiggle of snakes and ladders to be 0 and 90 deg respectively. Since, as May and Hess suggested, within the critical spacing, aligned gabors bind more strongly, gabors should bind more strongly with each other in a snake than in a ladder. This makes snakes more detectable than ladders.

Dakin and Baruch (2009) found that a snake contour's shape is easy to identify (2afc) when the distracter gabors immediately around it are oriented nearly perpendicular to the contour path but hard to identify when they are nearly parallel to it. Ladders are hard to identify in the presence of either perpendicular or parallel distracters. The stronger binding of aligned gabors predicts that, in the case of snakes, making the distracters nearly parallel will improve alignment and binding with the target gabors; whereas making the distracters nearly perpendicular will worsen alignment and binding with the target gabors. This will make the snake harder to detect among nearly parallel than among nearly perpendicular distracters. In the case of ladders, gabors on the path are more easily bound to grass gabors than to each other, making the ladder hard to detect.

Nugent, Kesawani, Woods, and Peli (2003) asked observers to detect a contour among random distracters (a snake in the grass). The same-sized stimulus was presented at various eccentricities. They found that contour detection falls with increasing eccentricity. This result too is explained by the finding that, within the critical spacing, aligned gabors are more strongly bound. With increasing eccentricity, the combining field grows, enclosing more and more grass within it. This leads to more spurious binding of target gabors to distracter gabors, thus making it harder to detect the contour. Or, as May and Hess (2007) put it, "If the snake elements are not completely collinear then, as the field size increases, there is an increased probability that an element of a curved snake will be more collinear with a distractor element."

Thus, the results of all these studies are explained by the single finding that, within critical spacing, aligned gabors bind more strongly than misaligned gabors. Binding target to target yields contour integration. Binding target to distracter disrupts contour detection. This same binding of target to flankers makes the target harder to identify, which is called crowding. This supports the suggestion by May and Hess (2007) that the association field that explains contour integration is none other than the combining field that explains crowding.

Conclusions

Alignment aids contour integration but impairs identification of a target among flankers. We find that the sensitivities to wiggle for these two very different tasks are equal, suggesting that the binding in the two phenomena is produced by the same or very similar mechanisms. It has previously been suggested that grouping among flankers reduces their crowding of the target. Instead, our review concludes that crowding is mediated by grouping of the flankers with the target and is unaffected by grouping of the flankers with each other.

Acknowledgments

References

Andriessen JJ, Bouma H. Eccentric vision: adverse interactions between line segments. Vision Research. 1976; 16 (1):71– 78. [PubMed: 1258390]

Appelle S. Perception and discrimination as a function of stimulus orientation: the "oblique effect" in man and animals. Psychological Bulletin. 1972; 78 (4):266– 278. [PubMed: 4562947]

Bouma H. Interaction effects in parafoveal letter recognition. Nature. 1970; 226:177– 178. [PubMed: 5437004]

Brainard DH. The Psychophysics Toolbox. Spatial Vision. 1997; 10(4):433– 436. [PubMed: 9176952]

Campbell FW, Kulikowski JJ, Levinson J. The effect of orientation on the visual resolution of gratings. Journal of Physiology. 1966; 187 (2):427– 436. [PubMed: 5972182]

Dakin SC, Baruch NJ. Context influences contour integration. Journal of Vision. 2009; 9(2):13, 1–13.

Dakin SC, Cass J, Greenwood JA, Bex PJ. Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. Journal of Vision. 2010; 10(10):14, 1–16. [PubMed: 20884479]

Ellis, WD. A Source Book of Gestalt Psychology. London: K. Paul, Trench, Trubner & Co; 1938.

Field DJ, Hayes A, Hess RF. Contour integration by the human visual system: Evidence for a local "Association Field". Vision Research. 1993; 33:173– 193. [PubMed: 8447091]

Field DJ, Hayes A, Hess RF. The roles of polarity and symmetry in the perceptual grouping of contour fragments. Spatial Vision. 2000; 13:51– 66. [PubMed: 10688229]

Geisler WS. Visual perception and the statistical properties of natural scenes. Annual Review of Psychology. 2008; 59:167– 92.

Hess RF, Dakin SC. Contour integration in the peripheral field. Vision Research. 1999; 39:947– 959. [PubMed: 10341947]

Hess RF, Field DJ. Integration of contours: new insights. Trends in Cognitive Science. 1999; 3:480– 486.

Intriligator J, Cavanagh P. The spatial resolution of visual attention. Cognitive Psychology. 2001; 43 (3):171– 216. [PubMed: 11689021]

Kooi FL, Toet A, Tripathy SP, Levi DM. The effect of similarity and duration on spatial interaction in peripheral vision. Spatial Vision. 1994; 8 (2):255–279. [PubMed: 7993878]

Kovacs I, Julesz B. A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. Proceedings of the National Academy of Sciences of the United States of America. 1993; 90:7495– 7497. [PubMed: 8356044]

Ledgeway T, Hess RF, Geisler WS. Grouping local orientation and direction signals to extract spatial contours: Empirical tests of "association field" models of contour integration. Vision Research. 2005; 45 (19):2511– 2522. [PubMed: 15890381]

Levi DM. Crowding – An essential bottleneck for object recognition: A mini-review. Vision Research. 2008; 48:635– 654. [PubMed: 18226828]

Levi DM, Hariharan S, Klein SA. Suppressive and facilitatory spatial interactions in peripheral vision: peripheral crowding is neither size invariant nor simple contrast masking. Journal of Vision. 2002; 2 (2):167– 177. [PubMed: 12678590]

Livne T, Sagi D. Configuration influence on crowding. Journal of Vision. 2007; 7(2):4, 1–12. [PubMed: 18217819]

Loffler G. Perception of contours and shapes: Low and intermediate stage mechanisms. Vision Research. 2008; 48:2106– 2127. [PubMed: 18502467]

Martelli M, Majaj NJ, Pelli DG. Are faces processed like words? A diagnostic test for recognition by parts. Journal of Vision. 2005; 5:58– 70. [PubMed: 15831067]

May KA, Hess RF. Ladder contours are undetectable in the periphery: A crowding effect? Journal of Vision. 2007; 7(13):9, 1–15. [PubMed: 17997637]

Nugent AK, Kesawani RN, Woods Rl, Peli E. Contour integration in peripheral vision reduces gradually with eccentricity. Vision Research. 2003; 43:2427– 2437. [PubMed: 12972393]

Palomares M, LaPutt MC, Pelli D. Crowding is unlike ordinary masking [Abstract]. Investigative Ophthalmology and Visual Science. 1999; (Suppl):40, S351. [Abstract].

Parkes L, Lund J, Angelucci A, Solomon J, Morgan M. Compulsory averaging of crowded orientation signals in human vision. Nature Neuroscience. 2001; 4 (7):739– 744.

Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vision. 1997; 10 (4):437– 442. [PubMed: 9176953]

Pelli DG, Majaj NJ, Raizman N, Christian CJ, Kim E, Palomares MC. Grouping in object recognition: The role of a Gestalt law in letter identification. Cognitive Neuropsychology. 2009; 26(1):36– 49. [PubMed: 19424881]

Pelli DG, Palomares M, Majaj N. Crowding is unlike ordinary masking: Distinguishing feature detection and integration. Journal of Vision. 2004; 4 (12):1136– 1169. [PubMed: 15669917]

Pelli DG, Tillman KA. The uncrowded window of object recognition. Nature Neuroscience. 2008; 11(10):1129–1135. [Supplement].

Petrov Y, Popple AV, McKee SP. Crowding and surround suppression: Not to be confused. Journal of Vision. 2007; 7(2):12, 1–9. [PubMed: 18217827]

Saarela TP, Sayim B, Westheimer G, Herzog MH. Global stimulus configuration modulates crowding. Journal of Vision. 2009; 9(2):5, 1–11. [PubMed: 19271915]

Saarela TP, Westheimer G, Herzog MH. The effect of spacing regularity on visual crowding. Journal of Vision. 2010; 10(10):17, 1–7. [PubMed: 20884482]

Smits JT, Vos PG, van Oeffelen MP. The perception of a dotted line in noise: a model of good continuation and some experimental results. Spatial Vision. 1985; 1 (2):163– 177. [PubMed: 3940057]

Stuart JA, Burian HM. A study of separation difficulty: Its relationship to visual acuity in normal and amblyopic eyes. American Journal of Ophthalmology. 1962; 53:471– 477. [PubMed: 13917936]

Toet A, Levi DM. The two-dimensional shape of spatial interaction zones in the parafovea. Vision Research. 1992; 32:1349– 1357. [PubMed: 1455707]

Watson AB, Pelli DG. QUEST: A Bayesian adaptive psychometric method. Perception and Psychophysics. 1983; 33:113– 120. [PubMed: 6844102]

Wertheimer M. Laws of organization in perceptual forms. Untersuchungen zur Lehre von der Gestalt, II. Psychologische Forschung. 1923; 4:301–350. Translated in Ellis (1938).

Westheimer G. Meridional anisotropy in visual processing: implications for the neural site of the oblique effect. Vision Research. 2003; 43 (22):2281– 2289. [PubMed: 12962986]

Wilkinson F, Wilson HR, Ellemberg D. Lateral interactions in peripherally viewed texture arrays. Journal of the Optical Society of America A. 1997; 14(9):2057– 2068.

Yeotikar NS, Khuu SK, Asper LJ, Suttle CM. Configuration specificity of crowding in peripheral vision. Vision Research. 2011; 51(11):1239– 1248. [PubMed: 21458478]
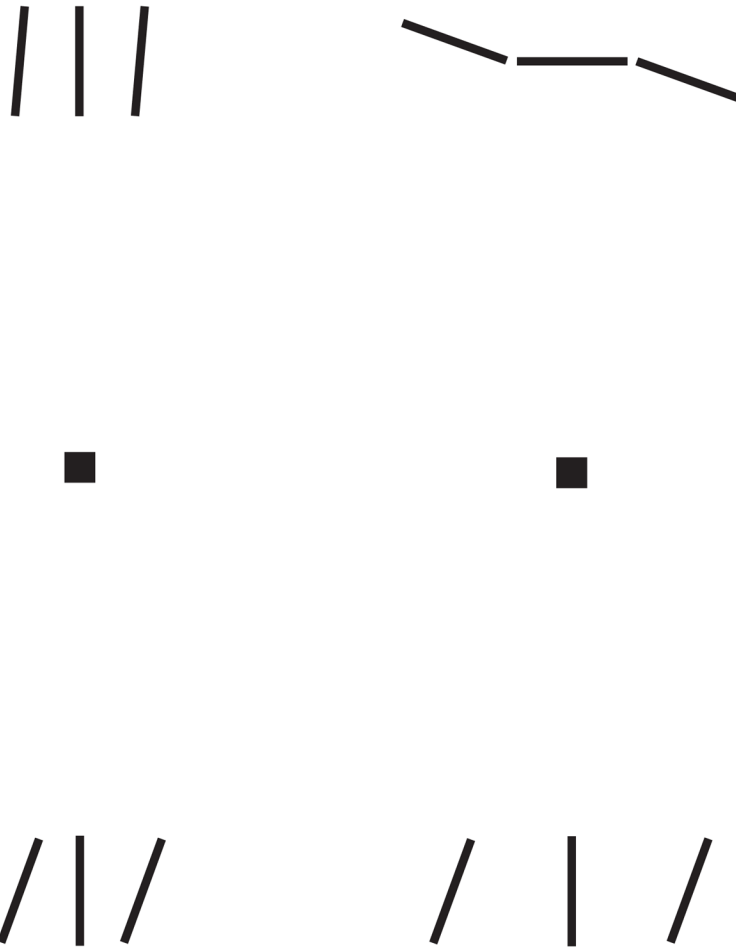
Figure 1. The effects of similarity and alignment in crowding
Each triplet consists of a target bar between two flanking bars. While fixating a square, judge the orientation of the target bar above or below fixation. **Left. The effect of similarity:** The upper and lower triplets in the left column differ only in the orientation of the flankers. While fixating on the left square, it is hard to tell that the upper target, which is similar to its flankers (+5 deg), is perfectly vertical but easy to tell that for the lower target, which is dissimilar to its flankers (+20 deg). **Right. The effect of alignment:** In the right column, the two triplets differ in alignment. The similarity of the flankers to the target is the same in both cases (+20 deg). While fixating on the right square, the upper target, which is aligned with its flankers, appears tilted anticlockwise from the horizontal whereas the lower target, which is misaligned with its flankers, appears perfectly vertical, as it is. The known differences in crowding between the upper and lower visual fields are too small to matter here; turning the page upside down has hardly any effect on these demonstrations.
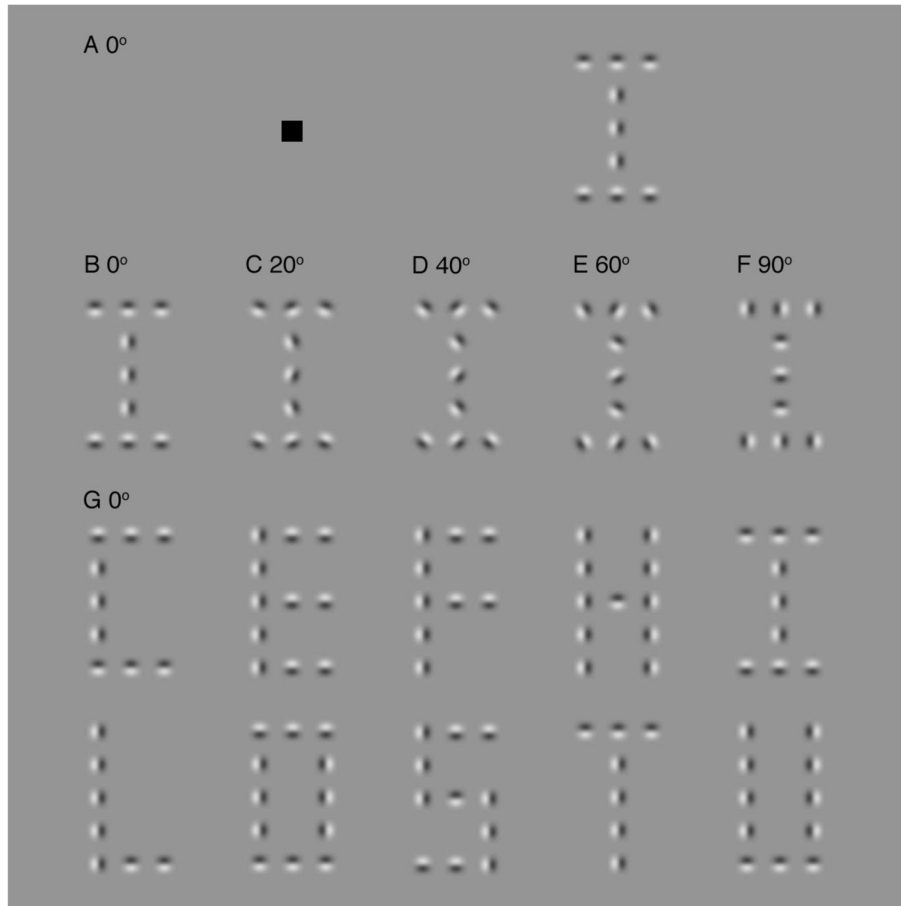
Figure 2. Snake letters
A: A "snake" letter made of gabors is presented 10 deg from fixation in the right visual field. A *gabor* is a grating patch (Eq. 1). The task is either to identify the letter or to report the phase of the central gabor (i.e., to indicate whether the light half of the gabor was on the right side or the left). **B–F:** Examples of letters (all "I") with wiggles of 0, 20, 40, 60 and 90 deg. **G:** All 10 letters, with zero wiggle, used in the contour integration task.
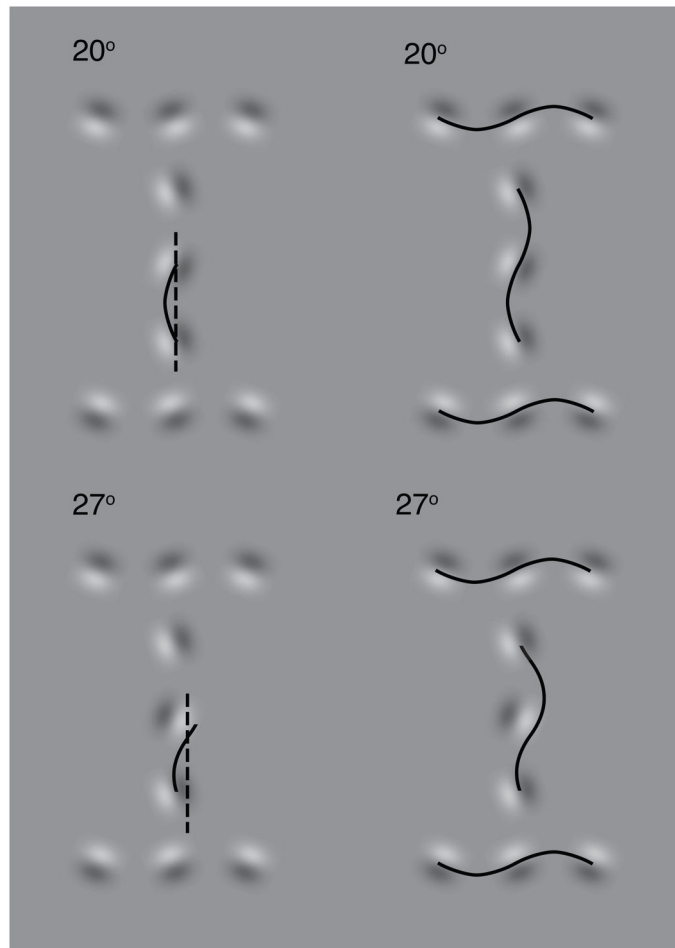
Figure 3. How to measure wiggle
The displayed value is wiggle. The orientation jitter is ±20°. The phase of the central target gabor is 0° in the upper row and 180° in the lower row. To measure a letter's wiggle, select a chain of gabors along a straight part of the original letter's path. Select a pair of adjacent gabors in that chain. Fit half a period of a sine wave, making tangential contact (at each end) with the dark-to-light transition of each gabor in that pair. The period of the sine fit is set to twice the center-to-center spacing of the gabors. The sine's position, orientation, amplitude, and phase are adjusted for best fit. The angle that the sine wave (solid) makes with its own axis (dashed) is that pair's wiggle. Repeat this for every pair of adjacent gabors in all straight chains in the letter. The average wiggle of all such gabor pairs is the letter's *wiggle*. The left column shows a single half-sine fit. The right column shows them all. Note that the sine waves are tangent to the dark-to-light transitions of the gabors, not the light-to-dark transitions. At the point of contact, the dark side of the transition is always to the right of the sine as the sin( ) argument increases. The sine fit need not pass through the center of the gabor and its axis need not be aligned with the letter path.
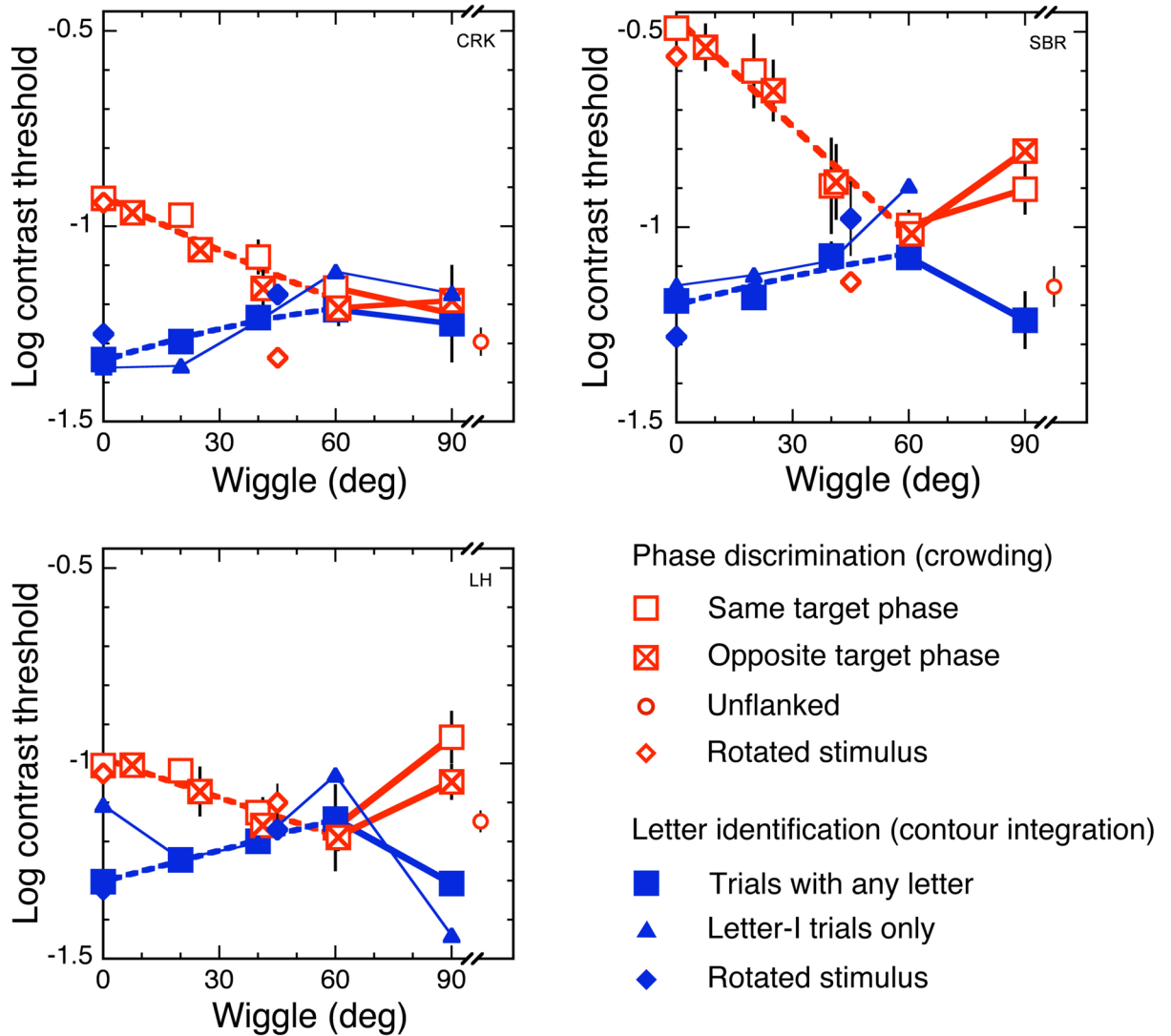
Figure 4. Three observers' results for both tasks

Increasing wiggle from 0 to 60 deg makes letter identification (solid symbols) harder and phase discrimination (open symbols, including the crossed squares) easier. In particular, contrast threshold for letter identification (solid squares) increases with wiggle, while contrast threshold for phase discrimination (open and crossed squares) of a single gabor embedded among other gabors decreases with wiggle. Phase discrimination data are split (post hoc) into same-target-phase (open squares) and opposite-target-phase (crossed squares) trials. Phase affects both wiggle and threshold in a consistent way, so that data from both phases trace out the same curve. Phase discrimination threshold for an unflanked gabor (small open circle, on the far right of each graph) serves as a baseline for the flanked phase thresholds. Letter identification is done with a 10-letter alphabet (solid squares), but analyzing just the letter-I trials yields very similar contrast thresholds (small solid triangles), showing that the letter differences ("I" vs. others) do not affect the measured thresholds. Contrast thresholds for snake letters rotated 45 deg clockwise from vertical (small diamonds) are similar to those for upright letters (squares). Error bars are mean ± standard error of log thresholds. All points, except thresholds for letter identification in letter-I trials, have error bars, but the error bars are invisible when smaller than the symbol. The effect of

wiggle up to 60 deg is monotonic. Beyond 60 deg, it varies among observers. In every observer, however, the effect of wiggle on each of the two tasks is a mirror image of the other. In the above plots, the data points in the 0 to 60 deg range of wiggle are left unconnected for unobstructed view of the quadratic fits (dashed). The data points beyond that range are connected by straight lines (solid).
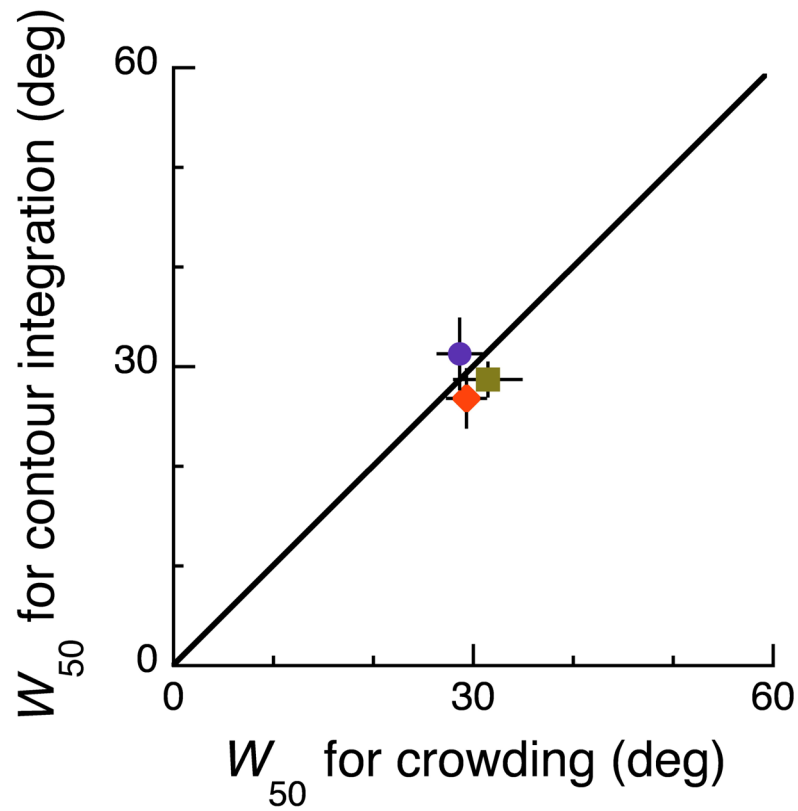
Figure 5. Scatter plot of wiggle thresholds
Each point is the wiggle threshold $W50$ in the contour-integration task versus that in the crowding task. As predicted, the wiggle threshold is about 30 deg and the data points all lie practically on the line of equality. For each observer, we obtain four independent wiggle thresholds in the letter identification task and plot the mean (observer SBR orange diamond; LH green square; CRK purple circle). For each observer, we obtain two independent wiggle thresholds for the phase discrimination task and plot the mean. Error bars are mean ± standard error.
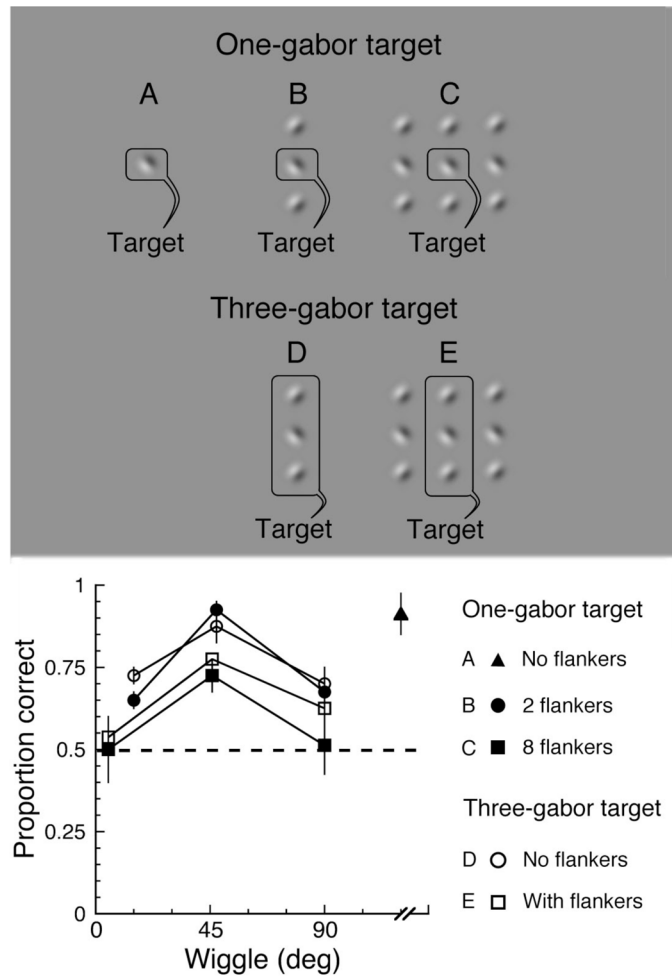
Figure 6. Crowding is unaffected by the spatial extent of the task-related information
The top panel displays the stimuli, which are presented in the periphery, 10 deg to the right of fixation. **A:** a single unflanked target gabor. **B:** a target gabor flanked by gabors, above and below, forming a vertical line. **C:** a target gabor flanked by eight gabors, forming three vertical lines. **D:** three target gabors in a vertical line, unflanked by other gabors. **E:** three target gabors flanked by two vertical lines of gabors. On a given trial, all gabors in the stimulus have one of three orientations relative to the vertical, 0, 45 or 90 deg, with the direction of tilt alternating between adjacent gabors. We add a small orientation jitter to each gabor, randomly chosen from a uniform distribution between -5 and 5 deg. The phase of each target gabor is random ($\psi = 0$ or 180 deg) on each trial. The non-target gabors have a phase of 0 deg. The wiggle is measured by the procedure described in Methods. For each configuration and wiggle, we present 40 trials. The stimulus is displayed for 150 ms. In both one-gabor-target and three-gabor-target conditions, after each trial, two response alternatives are presented on the screen. One is identical to the target and the other is a foil, identical to the target except that the phase of each gabor is random (0 or 180 deg), with the proviso that the foil is never wholly identical to the target. (The phase of more than one gabor can differ between target and foil when the target consists of several gabors.) The observer is asked to pick the alternative that matches the target. The bottom panel plots the (average) results of two observers. Performance in all conditions peaks at intermediate wiggles and declines at both higher and lower wiggles, just as observed in the main experiment. Chance performance (50%) is indicated by the dashed line. Error bars are mean ± standard error.

To test whether there is an effect of the area occupied by the task-related information, we compare performance across the two main conditions (single- versus three-gabor target) in the cases that display the same number of gabors on the screen, i.e. B versus D (solid versus open circles) and C versus E (solid versus open squares). At each wiggle, performance is similar in these two conditions, showing no effect of tripling the number of gabors (and area) containing task-relevant information.
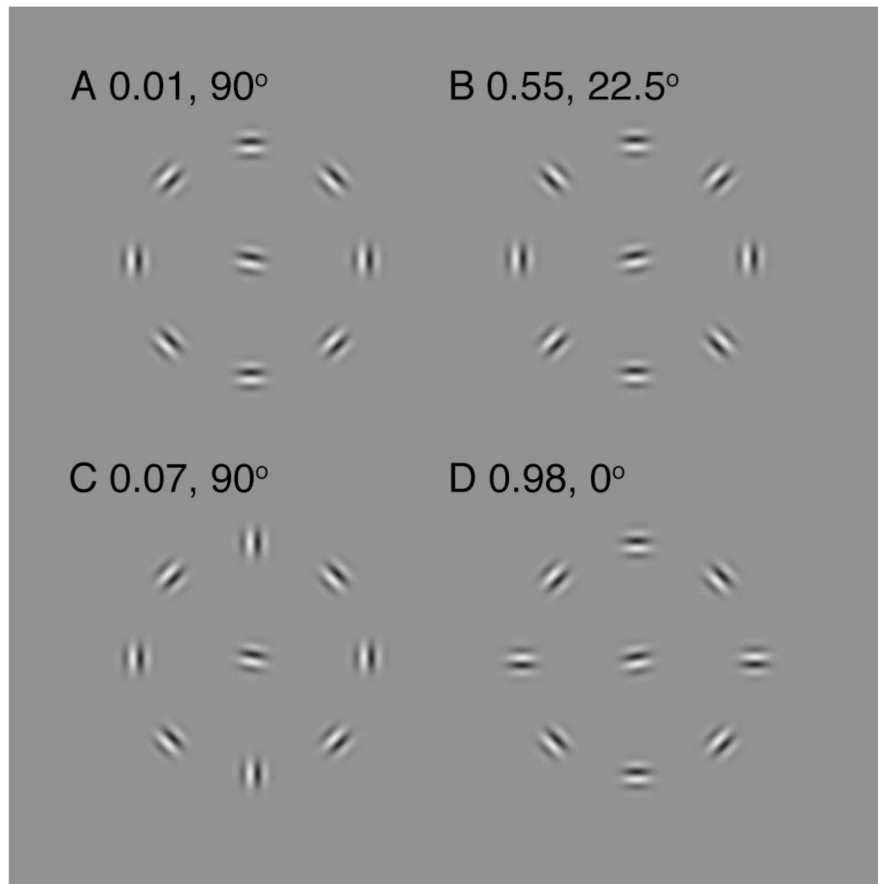
Figure 7. Livne and Sagi (2007) stimuli (A, B, C, D) with log threshold elevation (measured by them) and wiggle (measured by us)

Each stimulus consists of eight flanker gabors arranged in a circle around a target gabor. While fixating to the left or right of the stimulus, the task is to determine whether the central target is tilted clockwise or counterclockwise from horizontal. The configuration of the flankers was manipulated: **A.** all flankers group into a contour; **B.** the contour was interrupted at every other location; **C.** contour interrupted only at top and bottom; **D.** contour interrupted only at left and right. In these stimuli the target lies on a straight path with each pair of diametrically opposed gabors. We compute the wiggle of each of these gabor triplets in each of these configurations as a predictor of how well the target will bind with that pair of flankers. For each stimulus, we list the log threshold elevations and the lowest of the 4 wiggles. The two numbers reveal a reciprocal relation between log threshold elevation and minimum wiggle.

Livne and Sagi proposed that grouping flankers into a contour reduces crowding. As predicted, configuration A does not show any crowding, and B does. We averaged the observers' log threshold elevations for these configurations: with low contrast stimuli, average threshold elevation for A is 0.01±0.03 log units, and for B is 0.55±0.04 log units (n=2); threshold elevations for high contrast stimuli were similar (A 0.06±0.1, B 0.72±0.17; n=3). Their hypothesis predicts no difference in performance between configurations C and D. Our hypothesis is that the target is crowded whenever the target and flankers are grouped (aligned), regardless of grouping among flankers. Like theirs, our hypothesis correctly predicts much less crowding in A than in B. However, unlike theirs, our hypothesis predicts no crowding in C and strong crowding in D. In fact, this is what Livne and Sagi found (threshold elevation C 0.07±0.04, D 0.98±0.12; n=1). Furthermore, configuration D

(although it is interrupted in only two locations) showed the highest threshold increase of all configurations. Our hypothesis predicts this as well, since the alignment of the target with flankers is best in this configuration.