Temporal properties of the polarity advantage effect in crowding

Ramakrishna Chakravarthi
Vision Sciences Laboratory, Harvard University, Cambridge, MA, USA

Patrick Cavanagh
Vision Sciences Laboratory, Harvard University, Cambridge, MA, USA, & Département de Psychologie, Université de Paris 5, Paris, France

If the target in a crowding display differs from the distracters in its contrast polarity, the extent of crowding is reduced compared to the condition where all the elements in the display have the same polarity. In Experiment 1, we test the temporal properties of this *polarity advantage* by reversing the contrast of the target and flankers at four frequencies between 2 and 15 Hz. In the same polarity condition, target and distracters were all white in one frame but all black in the next. In the opposite polarity condition, the target was white and distracters black in one frame and all reversed in the next frame. Less crowding was seen for the opposite polarity condition at lower frequencies, but this advantage disappeared at 7.5 Hz and higher frequencies. In Experiment 2, we test whether this result can be explained by lateral masking, using a display that matched the crowding configuration. Lateral masking did not exhibit a polarity advantage at any frequency. Hence, the polarity advantage in crowding, and its loss at 6–8 Hz, cannot be attributed to lateral masking. It is known that attention has a coarse temporal resolution (6–8 Hz). The findings of this study suggest a role for attention in crowding, as opposed to low-level mechanisms like lateral masking.

Keywords: crowding, polarity advantage, temporal properties, attention, lateral masking

Citation: Chakravarthi, R., & Cavanagh, P. (2007). Temporal properties of the polarity advantage effect in crowding. *Journal of Vision, 7*(2):11, 1–13, http://journalofvision.org/7/2/11/, doi:10.1167/7.2.11.

Introduction

A target that can be easily identified on its own becomes much harder to identify when surrounded by flankers (Andriessen & Bouma, 1976; Bouma, 1970, 1973; Westheimer, Shimamura, & McKee, 1976). This effect of crowding increases with the similarity of the target and the distracters (Andriessen & Bouma, 1976; Banks & Prinzmetal, 1976; Kooi, Toet, Tripathy, & Levi, 1994) but cannot be erased by increasing the viewing time (Townsend, Taylor, & Brown, 1971). Outside the fovea, the critical spacing between the target and flankers at which identification drops to threshold level is surprisingly large, typically equal to about 1/3 to 1/2 the eccentricity. At the fovea it is much smaller, 1/10th of a degree or less (Toet & Levi, 1992; Wolford & Chambers, 1984).

We can divide the several theories that seek to explain crowding into two broad classes: *low level* and *high level*. An example of a low-level theory is lateral masking. Lateral masking is said to be due to feature interactions or competition for access to a limited number of feature detectors (Townsend et al., 1971; Wolford & Chambers, 1984). Target responses are degraded or inhibited by low-level masking from neighboring activity of the distracters at the level of retina, lateral geniculate, or primary visual cortex.

High-level theories, on the other hand, propose that the interaction between the target and flankers occurs at some later stage in the visual processing stream. For example, the "compulsory pooling of signals" hypothesis (Levi, Hariharan, & Klein, 2002; Levi, Klein, & Hariharan, 2002; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli, Palomares, & Majaj, 2004; Strasburger, 2005) suggests that the information collected to identify a target must be pooled over some area and when that area includes flankers, the target signal is no longer available/sufficient for identification—it might be averaged with the distracter signals or might be rendered inaccessible. Following a stage of feature detection, this pooling is said to occur at a second—"feature integration"—stage, the former being unaffected in crowding. The hypothesis that crowding is due to poor resolution of attention (Intriligator & Cavanagh, 2001) makes a similar point. The attention resolution theory specifies that the pooling mechanism is the selection step of visual attention. Compulsory pooling occurs because attention has a minimum available size of the selection region at a given eccentricity and that minimum size is quite a bit larger than the smallest resolvable detail.

doi: 10.1167/7.2.11 Received November 8, 2005; published March 27, 2007 ISSN 1534-7362 © ARVO

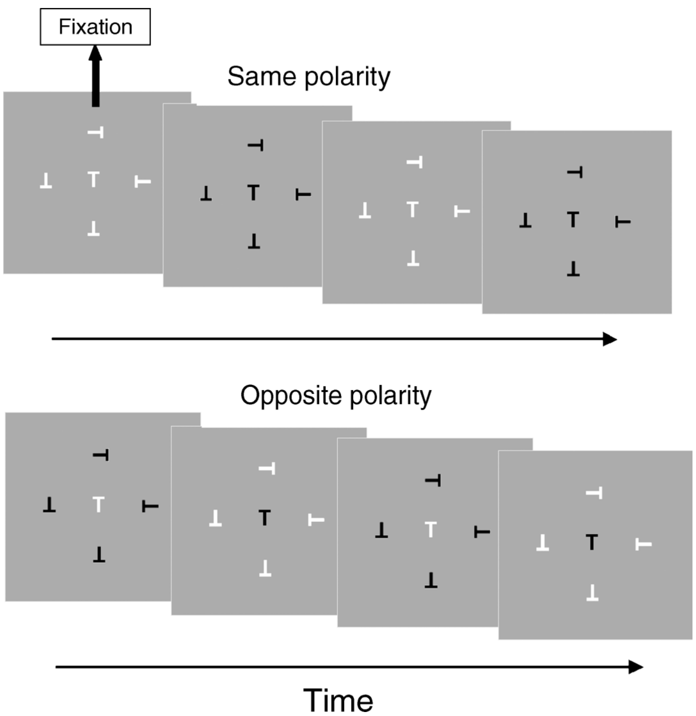

Figure 1. Individual segments of the display. Top row: same polarity condition. All items have the same contrast polarity at a given instant. The elements in the display alternate in polarity (all items simultaneously) without changing orientation or position in a given trial with gray intervals interspersed in Experiment 1A (not shown), but with no gray intervals in Experiment 1B. The gray segments are the same as the above but with no elements in the display except the fixation point. Bottom row: opposite polarity condition. The stimulus setup is the same but the central target has a different polarity relative to the distracters in each segment.

Two studies have shown that crowding is not simply the loss of the target signal at an early level in the visual system. He, Cavanagh, and Intriligator (1996) showed that crowded grating patches could induce orientation-specific aftereffects although observers were unable to report the orientation of the patches. Although the targets were crowded sufficiently to fall below threshold for identification, they were registered up to at least a level where orientation analysis does occur—at least V1 in the visual cortex—as evidenced by the orientation-specific aftereffect that remains significant but somewhat reduced in strength under crowded conditions (Blake, Tadin, Sobel, Raissian, & Chong, 2006). Parkes et al. (2001) showed that the orientation of a target in a crowded array of gabor patches could not be reported but, nevertheless, the estimate of the average orientation of the gabors in the array was directly influenced by the orientation of the target. These authors concluded that although the individual identities in a crowded array are blocked from awareness, they nonetheless do get through to higher levels, albeit in the form of a contribution to the array's overall texture.

We will further discriminate between low- and high-level theories by examining the temporal properties of the *polarity advantage* reported for crowding (Kooi et al., 1994). The strength of crowding is known to be strongly affected by the degree of similarity between target and flanks. Kooi et al. (1994) found that crowding was reduced when the flanks were of a different contrast polarity, shape, or (for some subjects) color compared to the target. In this paper, we will focus on the polarity advantage: the finding that distracters crowd a target of the same polarity much more effectively than a target of opposite polarity.

On its own, the polarity advantage can be explained by either low- or high-level theories of crowding. For example, lateral masking, as tested with a metacontrast stimulus, has been shown to be polarity specific (Becker & Anstis, 2004). So, if crowding is due to lateral masking and if this lateral masking has the properties seen in the metacontrast results of Becker and Anstis (2004), crowding may be reduced when the target and distracters have opposite polarities. In high-level models, for example, in the attention-based model, crowding occurs following the operation of selective attention. In this case, when the target item is a different polarity from the distracters, it can be selected based on its polarity and the distracters can be filtered out, again reducing crowding.

Although both approaches could support a polarity advantage, the two levels have very different temporal properties. The temporal properties of high-level processes, specifically attention, have been studied using several different approaches. In particular, the property of attention that is of interest here is selection: the ability to individuate an item from those that surround it spatially (spatial resolution) or those that precede and follow it temporally (temporal resolution). If crowding is a failure of spatial resolution of attention (He et al., 1996; Intriligator & Cavanagh, 2001), then polarity will influence the spatial selection in our alternating contrast stimulus (Figure 1) only when the rate of alternation falls within the temporal resolution of selection. At rates too fast for individual selection of the light and dark phases, target and distracters will no longer be experienced as having different polarities. They will all appear as flickering items, and polarity-based filtering can no longer provide attention with privileged access to the target. The loss of temporal individuation for separate phases of an alternating stimulus has been referred to as Gestalt flicker fusion (van de Grind, Gruesser, & Lunkenheimer, 1973) and it marks the change from the subjective experience of the individual light and dark phases to the perception of flicker. It occurs at a temporal rate far lower than the flicker fusion rate.

The maximum rate at which this polarity individuation can be maintained has been measured in a variety of motion, apparent motion, tracking, and detection tasks (Battelli et al., 2001; Battelli, Cavanagh, Martini, & Barton, 2003; Rogers-Ramachandran & Ramachandran, 1991, 1998; Verstraten, Cavanagh, & LaBianca, 2000). These experiments imply a limit for the temporal resolution for selection of around 6–8 Hz (60–80 ms per frame) and so suggest a similar temporal limit for the effectiveness of the polarity advantage in crowding. Other rapidly alternating displays that require identification of individual stimuli (RSVP studies where each new stimulus masks the previous one) also indicate temporal properties in this range. Critical rates are often 10 to 14 images per second (Eriksen & Collins, 1969; Lawrence, 1971; Potter, 1975; Raymond, Shapiro, & Arnell 1992), indicating a minimum time for accessing an individual frame of about 80 ms. Other high-level models of crowding may also invoke slow temporal response properties for crowding and the polarity advantage, but the temporal properties of such candidates (such as "compulsory pooling of signals") have not yet been investigated.

Several other methods have been used to assess the time course of attentional processes as well, generally attributing a slower rate than the one determined for selection—usually on the order of 2–4 Hz. However, many of these processes—attentional blink (Raymond et al., 1992; Shapiro, Raymond, & Arnell, 1994), attentional gating (Reeves & Sperling, 1986; Sperling & Weichselgartner, 1995), spatial cueing (Cheal & Lyon, 1991; Nakayama & Mackeben, 1989), and rate of binding (Holcombe & Cavanagh, 2001)—require not just selecting the target (which is the process of interest to us) but also disengaging from an object, engaging another object, identifying an object, interpreting (central) cues, binding of spatially separated features, and/or in some cases invoking other higher mental process such as working memory. This explains the extremely slow time course for attention obtained in these studies.

In contrast to the temporal resolution of selective attention, lateral masking should operate up to much higher frequencies—at least up to 30 Hz: Hess and Snowden (1992) and Snowden and Hess (1992) showed that ordinary masking operates at high frequencies. The temporal properties of lateral masking, as an underlying mechanism of crowding, is only indirectly related to the existing evidence for polarity effects in masking at different temporal frequencies. For example, center–surround interactions have been shown to be polarity sensitive (contrast sensitivity is higher if the test area alternates out of phase with its surround than in the in-phase configuration; see Experiment 3 of Spehar & Zaidi, 1997), but this was tested only at low temporal frequencies (<2 Hz). Another study, similar to Spehar and Zaidi (1997) but testing at higher frequencies (4–20 Hz), reported that the perception of the strength of a flicker (and hence the ability to detect it) is strongly modulated by the phase relation between the target and its surround (Kremers, Kozyrev, Silveira, & Kilavik, 2004)—an out-of-phase center–surround being perceived as a stronger flicker than an in-phase one. A flickering target should therefore be easier to detect if surrounded by distracters of opposite polarity (flickering 180° out of phase). In this case, the contrast difference between it and its distracters would be greater than that in the same polarity condition and this can be expected to hold up to higher frequencies.

It must be noted that the lateral masking studies examined above presented stimuli at the fovea, and the spacing between the test and its surround was either absent or negligible. Hence, although this evidence is suggestive, a direct determination of the temporal properties of lateral masking has not yet been reported, particularly in the periphery with a stimulus configuration matched to that of a crowding display. This will be specifically tested in Experiment 2. If the pattern of polarity advantage, if any, for lateral masking differs from that for crowding, it will argue against a low-level lateral masking explanation for the polarity advantage.

Experiment 1

The differences between the temporal properties of attention and masking lead to very different predictions for the effect of varying alternation rate on the polarity advantage in crowding. In this experiment, we constructed a dynamic crowding display to test the predictions of the two models. We presented a standard crowding array of a central target flanked by four distracters (see top row, Figure 1) flickering at different frequencies with the polarity of the array reversing on every segment. There were two conditions—*same polarity* and *opposite polarity*. In the *same polarity* condition, all the items on display had the same contrast polarity with respect to the background and reversed their polarity on alternate segments. In the *opposite polarity* condition, the polarity of the target in each segment was the opposite of that of the distracters (second row, Figure 1), and all items reversed contrast on each segment so that the target retained its opposite polarity throughout.

How does this display discriminate between low- and high-level models? For the low-level model, we assume that the lateral masking is effective and polarity specific up to very high rates. Clearly, at very high rates of alternation the white and black versions of each letter will fuse into the same gray as the background (as the white and black letters are equidistant from the background in terms of contrast) and the task will become impossible. Nevertheless, a polarity advantage is expected at all rates up to the point where the display is no longer visible.

The expectation is very different for the high-level model. At frequencies below 6–8 Hz, we again expect to see the polarity advantage as each segment should be seen individually and the polarity differences or similarity should be effective. Beyond this limit, however, the white and black versions of each letter will no longer be experienced as individual white and black letters. They will appear as flickering letters on a gray background and

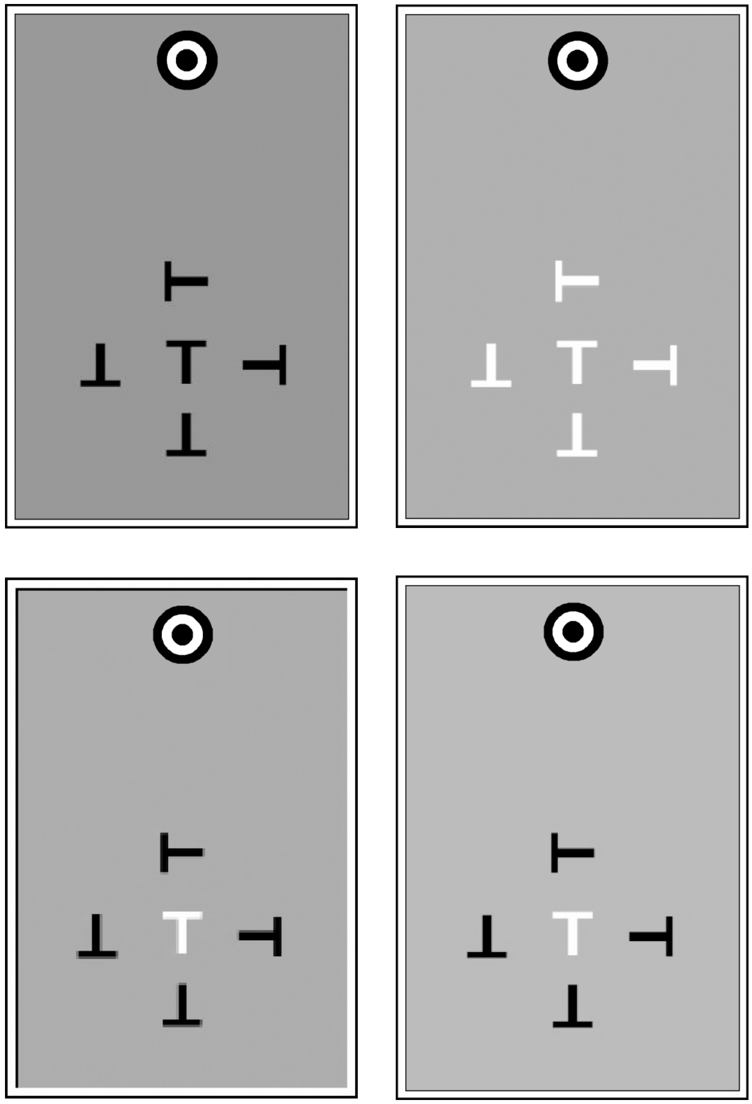

Figure 2. Movies of the crowding displays. Top left: same polarity display at 3 Hz; top right: same polarity at 12 Hz; bottom left: opposite polarity at 3 Hz; and bottom right: opposite polarity at 12 Hz. The movies illustrate the stimuli used in the experiment, but because of the timing limitations of the movie format they cannot provide an exact match to the experimental timing. All the movies are repeating sequences at a nominal 24 fps and each has just two frames with the target and distracters where the polarities switch between frames. At the fast rate, there are no intervening blank frames between target frames, giving a 12-Hz rate for the alternation of each cycle of two frames. For the slow rate, the target frames are separated by three blank frames for a total of eight frames per cycle and a 3-Hz rate. Click on each image to view the corresponding movie.

not as, for example, clearly visible white items surrounded by black, thus hampering any filtering mechanism that might provide attention with a privileged access to the target. For a high-level locus of crowding, one that follows the temporal limits of selection by attention, the effect of polarity on crowding should disappear beyond this rate.

Experiment 1A: Methods

Subjects

Four experienced subjects, aged 27–32 years, with normal or corrected to normal vision took part in this experiment.

Materials and stimuli

The basic setup involved production and display of stimuli by Vision Shell stimulus production software on an Apple Power Macintosh and 18 in. (14.5 × 11 in.) Sony monitor. The targets and distracters were gray scale *T*'s in four possible orientations (upright, upside down, rotated left, and rotated right) presented on a uniform gray background of luminance 34.2 Cd/m^2 (coordinates: $x = 0.307$, $y = 0.313$). At a viewing distance of 57 cm, the two bars that made up the *T*'s were 1.0° of visual angle in size. The contrast for the dark and light *T*'s was fixed at 20% of the background to ensure above threshold visibility at all frequencies. The eccentricity of the target

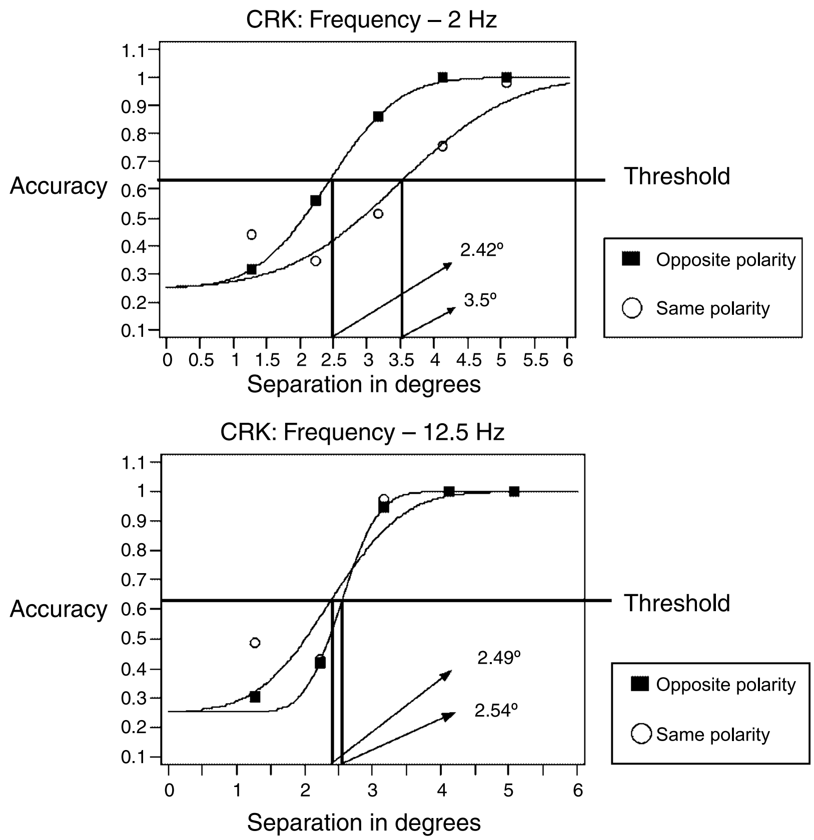

Figure 3. Determination of critical spacing. These graphs show psychophysical curves of performance accuracy plotted against the distance between the target and the distracters for the two conditions—same polarity (open circles) and opposite polarity (filled squares)—at two different frequencies in one subject. The critical spacing is the separation at which performance is at threshold (62.5% accuracy).

was fixed at 7.5° in the lower visual field. The distracters were presented at the ends of an imaginary cross centered on the target (see Figures 1 and 2). The distance between the target and the distracters was varied with trials, with the distracters being equidistant from the target in any given trial.

The displays alternated at one of four different temporal frequencies—2 Hz, 3.75 Hz, 7.5 Hz, or 12.5 Hz. Each cycle consisted of two frames containing the arrays of *T*'s, each followed by an interval of background gray. Thus, there were four segments to each cycle—the first frame, an interframe gray field (no elements on display), the second frame where all the polarities of the *T*'s of the first frame are reversed, and another interframe gray field (the interframe gray fields are not shown in Figure 1). The fixation spot was present in all segments. The duration of presentation of the crowding display frames (first and second) was fixed at two monitor refresh frames each (~27 ms). The duration of the gray fields varied with frequency, being shorter for higher frequencies (Figure 2). The cycles were run until the subject responded. This procedure was adopted in the light of the finding that a change in display duration affects the effective crowding distance (Tripathy & Cavanagh, 2002). We do not know, for the repeating frames that we use, whether the critical duration parameter is the individual frame duration or the total duration. In this first experiment, we keep the individual durations fixed, presenting the stimuli for the same duration (~27 ms) per stimulus frame in all conditions, in which case the total duration of stimulus exposure increases as the frame rate increases. In the following experiment (Experiment 1B), we use the opposite approach and keep the total duration fixed while letting the stimulus duration on individual frames decrease as frequency increases.

Procedure

There were two conditions. The same polarity condition where all the *T*'s in the display had the same polarity, and the opposite polarity condition where the target had a different polarity from the distracters at all times.

The method of constant stimuli was used. The two conditions, the four frequencies, and the varying target to distracter spacings were presented in a pseudorandom manner within each session. Feedback was provided. The task was to report the orientation of the central target by means of a key press (one of four possible responses).

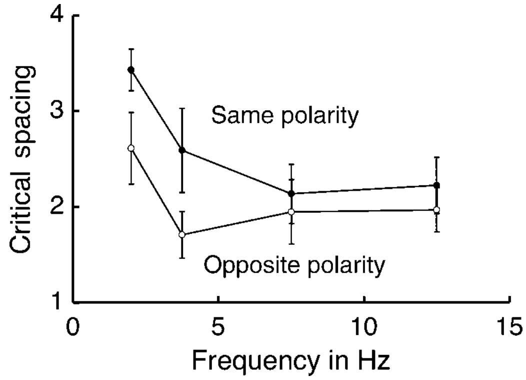

Figure 4. Results of Experiment 1A: The graph plots the critical spacing at various frequencies for two conditions—same and opposite polarities. The graph shows the average of four subjects. The *y* axis represents critical spacing measured in degrees. The same polarity curve (filled circles) and the opposite polarity curve (open circles) are separated at lower frequencies but converge at around 7.5 Hz. Error bars represent one standard error of the mean.

Accuracy was measured and evaluated for each frequency as a function of distance between the target and the surrounding distracters. Finally, critical spacing, the distance between the target and distracters at which performance is at threshold, was obtained from the psychophysical curves (probit function) of accuracy plotted as a function of distance for each frequency within each condition separately (see Figure 3). Threshold (62.5%) here was taken as midway between chance (25%) and maximal performance (100%). The critical spacing in the two conditions (same and opposite polarity) was then plotted as a function of frequency (Figure 4). The two curves were compared specifically to note the point of convergence—that is, the frequency at which there was no longer any significant polarity advantage.

Results

A repeated measures two-way ANOVA was conducted with frequency and condition as the factors and the critical spacing as the dependent variable.

Averaged across frequencies, there was less crowding in the opposite polarity condition than in the same polarity condition [distracters could be closer for threshold performance: average critical spacing 2.1° for opposite polarity vs. 2.6° for same polarity, $F(1, 3) = 44.61$, $p < .01$] consistent with earlier findings using nonalternating displays (Kooi et al., 1994).

Averaged over same and opposite polarity conditions, there was less crowding at higher temporal frequencies, $F(3, 9) = 14.12$, $p = .001$. This reduction of crowding with increasing frequency may be a result of the increase in the number of times a given stimulus is seen per unit time at higher frequencies compared to lower frequencies (the duration of each individual presentations is fixed at ∼27 ms for all frequencies. The blank intervals vary with frequency and thus one sees a larger quantity of stimulus per unit time at higher frequencies than at lower ones). The increase in total stimulus presentation per unit time across the repetitions may, like an increase in the duration of a single presentation (Tripathy & Cavanagh, 2002), decrease the radius over which crowding occurs. Another explanation might be based on the fact that alternating target and distracters become flicker-defined beyond 6–8 Hz rather than contrast-defined (Battelli et al., 2003; Rogers-Ramachandran & Ramachandran, 1998). It might be that crowding of flicker-defined objects differs from that of contrast-defined objects.

The interaction between condition and frequency was not significant, $F(3, 9) = 2.72$, $p > .1$, but the planned contrast testing our hypothesis of polarity advantage at low but not high frequencies was significant, $F(1, 24) = 7.32$, $p < .025$. Specifically, the critical spacing differed significantly between the two polarity conditions at the two lower frequencies but not at the two higher frequencies: Subjects showed less crowding in the opposite polarity condition than in the same polarity condition at low frequencies, as expected based on either the high-level or the low-level model, but not at higher frequencies. That is, the curves converged at around 7.5 Hz (see Figure 4), thus corroborating the high-level theories.

A pilot study ($n = 2$) was also run to check for the effect of eye movements, if any. Here, the polarity advantage effect with *static* images (no flicker) was tested in two blocks: The crowding stimulus was presented (i) randomly, in each trial, in either the lower or the upper visual field and (ii) always in the same (lower) visual field. The display duration was fixed at 125 ms, thus not allowing enough time for any eye movements. Both blocks showed a similar, strong polarity advantage—thus suggesting that the polarity advantage is not dependent on either eye movements or knowledge of the test location.

Experiment 1B: Methods

The reduction in crowding with increasing frequency seen in the previous experiment might be due to a greater total duration of stimulus presentation at higher frequencies. To control for this factor, we conducted a second experiment in which it was kept constant across all conditions. This was implemented by eliminating the interframe gray fields between stimulus frames and manipulating the frequencies by changing the duration of the stimulus frames.

Subjects

Four experienced subjects, aged 27–32 years, with normal or corrected to normal vision took part in this experiment. One of the subjects in this experiment (the first author) had also participated in Experiment 1A. The rest were naïve to the purpose of the experiment.

Materials and stimuli

The materials and stimuli were the same as in Experiment 1A except for the following changes. The eccentricity of the target was fixed at 9°. The sizes of the two component lines of the *T*'s used were 1.2° each. The displays alternated at one of four different temporal frequencies—2 Hz, 4 Hz, 7.5 Hz, or 15 Hz. Crucially, each cycle consisted of two frames containing the arrays of *T*'s, each of equal duration. The duration of the frames was varied to obtain various frequencies. No gray fields were inserted between frames as in Experiment 1A (Figure 1). This ensured that the signal per time remained constant at all frequencies.

Procedure

The procedure remained the same as in Experiment 1A.

Results

The data were analyzed in the same way as earlier. Figure 5 shows the averaged results for four subjects.

A repeated measures two-way ANOVA was conducted with frequency and polarity condition as the factors and critical spacing as the dependent variable.

There was a main effect of polarity, $F(1, 3) = 58.07$, $p = .005$, and also of frequency, $F(3, 9) = 4.22$, $p < .05$. That is, performance in the opposite polarity condition was significantly better than in the same polarity condition, and crowding at various frequencies was significantly different. Although the interaction of frequency and condition was not significant, $F(3, 9) = 2.43$, $p > .1$, the planned contrast that tested whether critical spacing across conditions would converge at 7.5 Hz was highly significant, $F(1, 24) = 13.96$, $p < .005$. This suggests that the advantage afforded by the polarity difference between the target and the distracter vanishes at around 7.5 Hz.

The averaged data in both Experiments 1A and 1B show that the polarity advantage is lost by 7.5 Hz. Further, for all four subjects in each experiment, there was a clear polarity advantage at the lowest frequency (2 Hz) and none at the highest frequency (15 Hz). There was some individual variation on the point (4 or 7.5 Hz) at which the advantage was lost but was seen to occur at the 7.5-Hz mark in most cases.

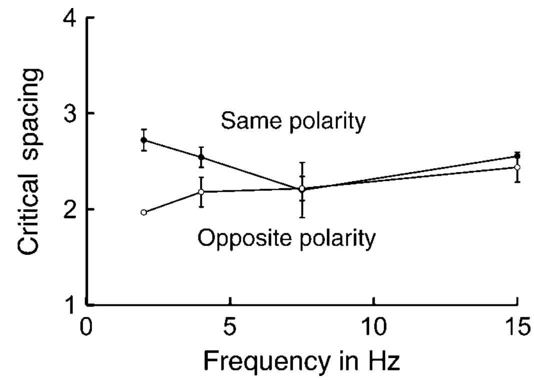

Figure 5. Results of Experiment 1B: The graph plots the critical spacing in degrees at various frequencies for the two conditions. The graph shows the average of four subjects. The same polarity curve (filled circles) and the opposite polarity curve (open circles) are separated at lower frequencies but converge at around 7.5 Hz. Error bars represent one standard error of the mean.

Experiment 2

Experiments 1A and 1B showed that the polarity advantage seen under crowded conditions is lost around 7.5 Hz, supporting the hypothesis that high-level processes, particularly attention, underlie crowding. However, the temporal properties of lateral masking in the periphery have not been explored directly so far. In this experiment, we test lateral masking as a function of polarity across the same range of frequencies tested with crowding, maintaining as well the spatial and temporal characteristics of the crowding displays. In this experiment, we use target detection (the subjects report the presence or absence of the central target) as a sensitive probe of lateral masking (as opposed to identification used in crowding).

This experiment consisted of two parts: one block with flankers present and one block with no flankers present. The comparison of performance in the two blocks would indicate whether any lateral masking occurs at the test frequencies. If yes, we can then specifically analyze the block with flankers, which consisted of two conditions—same polarity and opposite polarity. A comparison of performance between these two conditions would indicate the occurrence and extent of any polarity advantage in lateral masking.

Methods

Subjects

Four experienced subjects, aged 25–32 years, with normal or corrected to normal vision took part in the flankers-present block of this experiment and three subjects (two of whom had participated in the flankers-present block) took part in the flankers-absent block.

Materials and stimuli

The setup was made as similar as possible to the crowding display to determine whether lateral masking

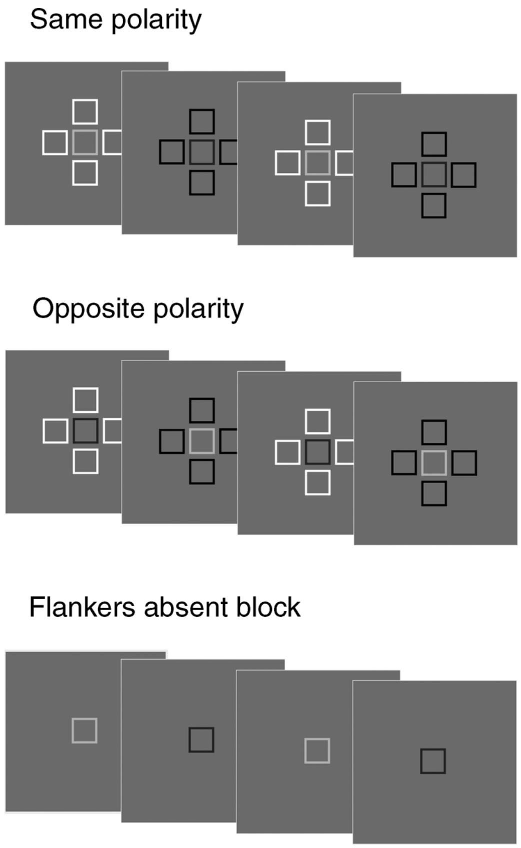

Figure 6. Lateral masking display—top and middle rows: flanker-present block. Top row: same polarity condition. All items have the same contrast polarity at a given instant. All items reversed polarity with every frame. As in Experiment 1B, there were no gray intervals between frames. Middle row: opposite polarity condition. The stimulus setup was the same but the central target had a different polarity relative to the distracters in each segment. Bottom row: flankers-absent block. In all conditions, the contrast of the target and the frequency were varied in each trial. In half of the trials, the target was present and in the other half absent. The subject had to indicate the presence or absence of the target.

operated under the same circumstances as crowding. The targets and distracters were gray scale squares. Squares instead of letters were chosen as that would allow closer apposition between the target and the distracters than the *T*'s while retaining several of the same features that make up the letters used in the earlier experiments. At a viewing distance of 57 cm, the squares were 1.4° of visual angle in size on each side. The eccentricity of the target was fixed at 9°. The target was presented at the center of an imaginary cross at whose ends distracters were presented (see Figure 6). The distance between the (centers of the) target and the distracters was fixed at 1.7°, so that the target and the distracters were close to but did not touch each other. The contrast for the dark and light phases of the distracters was fixed at 20% of the background, but the contrast of the target varied with trials. The displays alternated at one of four different temporal frequencies—2 Hz, 4 Hz, 7.5 Hz, or 15 Hz. Each cycle consisted of two frames containing the arrays of squares, each of equal duration. The duration of the frames was varied to obtain various frequencies. No gray fields were inserted between frames.

Procedure

The procedure was similar to that of the earlier experiments except for the following changes. There were two blocks: one with flankers present and the other with flankers absent. In the flankers-present block, there were two conditions: same polarity and opposite polarity as in the prior experiments. In this block, the two conditions, four frequencies, and the varying target contrasts were presented in a pseudorandom manner. In the flankers-absent block, there was no polarity condition and hence the four frequencies and the varying target contrasts were presented in a pseudorandom manner. Feedback was provided. In both blocks, the task was to report the presence or absence of the target square. The target was present in 50% of the trials and absent in the rest. Accuracy was measured and evaluated for each frequency as a function of target contrast. Finally, the threshold contrast was obtained from the psychophysical curves (probit) fitted for each frequency within each condition separately. Threshold (75%) contrast was taken as the contrast at which performance was midway between chance (50%) and the maximum (100%).

Results

The data were analyzed as for Experiments 1A and 1B. Figure 7 shows the averaged results for both blocks.

A two-way ANOVA was conducted with frequency and presence/absence of flankers as the factors and target contrast as the dependent variable. There was a main effect of flanker presence, $F(1, 20) = 91.9$, $p < .001$, and also of frequency, $F(3, 20) = 6.45$, $p < .005$, but the interaction was not significant, $F(3, 20) = 2.31$, $p > .1$. Thus, performance was significantly better if there were no flankers than in their presence. In other words, flankers could effectively mask targets. This was seen at all frequencies. The results also indicated that performance worsened with an increase in frequency. Further, a 1×4 ANOVA of the performance in the no-flanker block showed that there was an increase in contrast threshold with increase in frequency, $F(3, 8) = 15.1$, $p = .001$. This accords well with the literature on temporal contrast sensitivity functions: Contrast threshold increases with frequency (de Lange, 1958; Kelly, 1961).

A two-way repeated measures ANOVA with polarity and frequency as factors and target contrast as the dependent variable was performed on the data with flankers. There was no main effect of polarity, $F(1, 3) = 1.05$, $p > .1$, and the interaction was not significant, $F(3, 9) = 1.1$, $p > .1$, nor was the planned contrast comparing the polarity

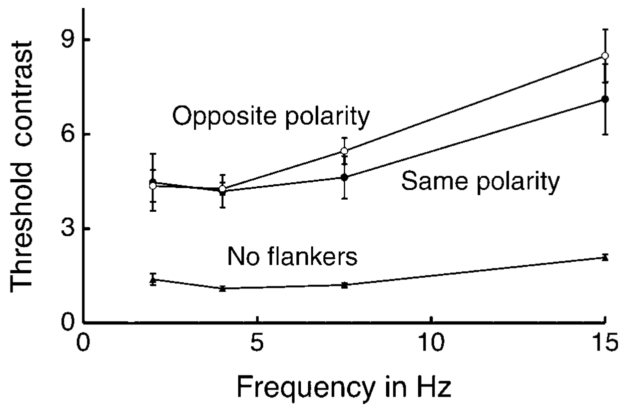

Figure 7. Results of Experiment 2: The graph plots the threshold contrast (with respect to the background) for target detection at various frequencies for each block: with flankers (circles) and without flankers (filled triangles). Performance in the two blocks is significantly different. The two conditions within the block with flankers: the same polarity curve (filled circles) and the opposite polarity curve (open circles) are statistically not different at any frequency. Error bars represent one *SEM*.

effect at the two low frequencies against that of the two high frequencies, $F(1, 32) = 0.0004$, $p > .2$. There was a main effect of frequency, $F(3, 9) = 11.1$, $p < .005$, as already determined by the earlier ANOVA. That is, contrary to our expectations, there was no polarity advantage at any frequency. In fact, the graph suggests a trend of *better* performance in the same polarity condition (lower threshold) compared to the opposite polarity condition. Critically, for our interpretation of Experiment 1, there was no advantage for opposite polarity in lateral masking across the range of frequencies tested. We conclude that the polarity advantage in crowding we found in Experiments 1A and 1B and the frequency dependence of this advantage cannot be attributed to lateral masking.

The absence of a polarity effect for lateral masking is a surprising finding, given the expectations based on the results from center–surround (Kremers et al., 2004; Spehar & Zaidi, 1997) and metacontrast masking (Becker & Anstis, 2004) experiments. Clearly, lateral masking in a configuration similar to that of crowding in peripheral vision is not polarity specific. Although there must be center–surround interactions that mediate the masking we do see, at the eccentricity, target and distracter size, spacing, and temporal rates that we use, these interactions are polarity independent. In other words, the differences between the configuration we use and those of the earlier experiments (Kremers et al., 2004; Spehar & Zaidi, 1997) might explain our contrary results. The primary difference was that in the earlier studies the target was presented in the fovea. Moreover, there was either a negligible or no gap between the target and the distracter, with the size of the distracter being several times larger than the target; also, the target and the surrounds were uniform circular fields and not letters. The configuration used in the current experiment more closely matches the traditional crowding setup and allows confidence in our conclusion that lateral masking under these conditions does not exhibit polarity specificity. Consequently, it can be argued that lateral masking does not play a significant role in the polarity advantage in crowding. Further testing would be necessary to identify the ranges of parameters in such a setup that do demonstrate polarity specificity and those that do not.

Discussion

The experiments here examined the temporal properties of the polarity advantage in crowding and lateral masking. Experiments 1A and 1B showed that a polarity difference helps in identifying a target under crowded conditions at low temporal frequencies. However, at higher frequencies this advantage is lost—this loss being manifest by around 7.5 Hz. We also found that no polarity advantage was detectable in lateral masking at any temporal frequency.

Thus, the results show that performance, in crowding but not in lateral masking, was better in the opposite polarity condition than in the same polarity condition up to frequencies of 7.5 Hz. Because attention has a temporal resolution of about 6–8 Hz (Verstraten et al., 2000), the polarity advantage appears consistent with the temporal limits of attention and not those of lateral masking. Our second experiment directly explored the temporal properties of lateral masking under conditions similar to those of our crowding experiments. We found that lateral masking can operate at frequencies as high as 15 Hz, but that a polarity advantage effect was not observed at any frequency. Lateral masking, under these conditions, does not seem to be polarity specific. Hence, lateral masking, as we have measured it, cannot explain the results found in Experiments 1A and 1B, thus strengthening our hypothesis that higher level processes such as attention underlie crowding.

It must be noted that a temporal resolution of attention of about 6–8 Hz does not imply that attention cannot select objects that are flickering at rates higher than this. Attention can and does select objects flickering at much higher frequencies. It only suggests that attention cannot individuate the phases (and consequently the identities) of the individual components that make up the flicker. That is, this limited resolution does not predict an increase in crowding distances at higher frequencies where the letters would be flicker-defined. It only predicts that attention will be unable to use any polarity information beyond its resolution frequency, and, if indeed attention underlies crowding, any advantage due to polarity differences in crowding will be lost, which is what we found.

Is the loss of polarity advantage then attributable to differences between contrast-defined (at low frequencies) and flicker-defined (at high frequencies) stimuli? Previous work supports the argument that alternating stimuli only act as flicker-defined stimuli when the alternating phases can no longer be individuated by attention (Rogers-Ramachandran & Ramachandran, 1991, 1998; van der Grind et al., 1973). This makes it reasonable to attribute the loss of the polarity advantage to the loss of the ability to individuate the separate polarities.

Additionally, in this first experiment, a strong main effect of temporal frequency on crowding was seen. An increase in frequency resulted in a decrease in crowding. As discussed earlier, this could have been a result of the stimulus presentation method adopted in that experiment—the amount of signal per unit time was correlated with frequency. In Experiment 1B, we controlled for this by keeping the stimulus presentation per unit time constant. Although, the effect of frequency was still significant, it was quite weak. Statistical analysis showed that the main contributor to the significance was the quadratic component of the curve (that is, more crowding at both low and high frequencies and less crowding at moderate frequencies) and the linear component was insignificant. This lends more support to the contention that the reduction in crowding at higher frequencies was due to an increase in signal per unit time at these frequencies. Further, and more importantly, this experiment apart from replicating the other finding of Experiment 1A—loss of polarity advantage by 7.5 Hz—demonstrated that this result was stronger under conditions of constant stimulus presentation per unit time.

The phenomenon of crowding has been extensively studied in the past (Bouma, 1970; Jacobs, 1979; Toet & Levi, 1992). Its similarity to certain low-level processes such as lateral masking, also known as "ordinary masking" or "simple contrast masking," led to the hypothesis that crowding resulted from interactions in the early visual system, particularly at the level of feature detectors (Wolford & Chambers, 1983, 1984; see also Pelli et al., 2004). Experimental evidence supporting this line of reasoning was in the form of discerning similarities between characteristics of crowding and masking. For example, it is known that crowding of a target is the greatest when the distracters are similar to it in terms of their spatial structure (Kooi et al., 1994) and spatial frequency (Chung, Levi, & Legge, 2001), which is what one would expect if it were due to lateral masking. Further, additional distracters increased crowding at closer spacings but not at larger spacings suggesting that there might be lateral interactions at closer spacings (Wolford & Chambers, 1983). In fact, the evidence for this hypothesis was thought to be so strong that the two terms—crowding and lateral masking—were used synonymously (for a recent analysis of the relationship between crowding and ordinary masking, including the terminology, see Huckauf & Heller, 2004). Although alternative explanations for crowding were proposed early on (response competition: Eriksen & Eriksen, 1974; higher level processes like attention and grouping: Banks, Larson, & Prinzmetal, 1979; Kahneman & Henik, 1977; Wolford & Chambers, 1983), it was generally accepted that featural interference was the basis of crowding.

Masking as an underlying explanation of crowding was called into question when several recent experiments (Chung et al., 2001; He et al., 1996; Levi, Hariharan, et al., 2002; Pelli et al., 2004; Strasburger, 2005; Wilkinson, Wilson, & Ellemberg, 1997) suggested the involvement of higher level processes. The hypotheses derived from such results were varied but they all had the underlying theme that lower level processes would not be sufficient to account for the phenomenon.

Further evidence for the dissimilarity between crowding and lateral masking is provided by Pelli et al. (2004), who show that key features of crowding differ from those of masking. In particular, the extent of crowding increases with eccentricity but not with signal size (Bouma, 1970; Levi, Klein, et al., 2002; Toet & Levi, 1992; Tripathy & Cavanagh, 2002). The opposite is true for masking. Masking seems to be independent of eccentricity but scales with signal size (Mullen & Losada, 1999). Masking affects identification as well as detection (Thomas, 1985a, 1985b), whereas crowding has little, if any, effect on detection but severely impairs identification (Pelli et al., 2004; Wilkinson et al., 1997). In masking, masks with low contrasts facilitate target detection, whereas at high contrasts they suppress the target according to a power function of their contrast (Legge & Foley, 1980). In crowding, there is no facilitation at any mask contrast, and the exponent of the power function is different from that of masking (Chung et al., 2001). Further, unlike in masking, this exponent varies as a function of spacing between the target and the distracters in a sigmoidal fashion, with a steep exponent at close spacings (Pelli et al., 2004). Masking is similar in the fovea and periphery whereas crowding in the periphery is qualitatively different from that in the fovea (Leat, Li, & Epp, 1999; Levi, Hariharan, et al., 2002; Levi, Klein, et al., 2002). Our findings can be added to this (growing) list of differences between crowding and lateral masking: Polarity-specific interaction is seen, at low temporal frequencies, in crowding but not at all in lateral masking when similar stimulus configurations are used.

Theories of crowding dealing with the architecture of attention emphasize the limited spatial and temporal resolution of its selection window (Cavanagh, 2004). Experiments (Intriligator & Cavanagh, 2001) have shown that the spatial resolution of attention is around half the eccentricity, very similar to the critical spacing of crowding. Similarly, the temporal resolution of attention is very poor: 6–8 Hz (Verstraten et al., 2000). He et al. (1996) suggested that the spatial resolution of attention was the source of the crowding phenomenon, and our experiments here suggest that the temporal resolution of attention restricts the effectiveness of the polarity advantage in crowding.

Further recent evidence of the interaction between attention and crowding comes from Poder (2005) who showed that attention when concentrated reduces crowding compared to when it is spread out. Huckauf and Heller (2004) reported a study that also suggests a fairly broad temporal window of integration for crowding. They presented distracters at various Stimulus Onset Asynchrony with respect to the target (−150 ms to +150 ms). They found that distracters interfered strongly with the identification of the target if presented anywhere in the window between 50 ms before the target and 100 ms after the target (Experiments 1A and 1B). This temporal window of 150 ms wherein information is integrated would imply a temporal resolution of about 6 Hz for crowding, which agrees with our findings and hypothesis.

Our experiments here relied on the loss of individuation of the black and white percepts in a flickering stimulus to manipulate the contrast polarity effect in crowding. The loss of temporal individuation for flicker has been referred to as Gestalt flicker fusion (Battelli et al., 2003; van de Grind et al., 1973). For example, Rogers-Ramachandran and Ramachandran (1991, 1998) presented an array of flickering discs in four quadrants with those in one quadrant flickering out of phase with the others. At flicker rates up to 7 Hz, subjects could identify the out-of-phase quadrant, whereas above 7 Hz performance was at chance, as all the discs appeared identical. Battelli et al. (2003) reported a similar result with a single patch out of phase with five others. If crowding arises at a level following this temporal limit of attention, there can be no further effect of contrast on crowding. This is what we found.

In summary, our results support the claim that crowding is not a low-level process (or at least not purely so) and has very coarse temporal limits.

Acknowledgments

This research was supported by grant NIH EY02958 to PC. We thank Laura Gibson for assistance in statistical analysis.

Commercial relationships: none.
Corresponding author: Ramakrishna Chakravarthi.
Email: chakrava@wjh.harvard.edu.
Address: Vision Sciences Laboratory, WJH, Harvard University, 33 Kirkland Street, Cambridge, MA 02143, USA.

References

Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research, 16,* 71–78. [PubMed]

Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics, 25,* 447–456. [PubMed]

Banks, W. P., & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception and Psychophysics, 19,* 361–367.

Battelli, L., Cavanagh, P., Intriligator, J., Tramo, M. J., Henaff, M. A., Michel, F., et al. (2001). Unilateral right parietal damage leads to bilateral deficit for high-level motion, *Neuron, 32,* 985–995. [PubMed] [Article]

Battelli, L., Cavanagh, P., Martini, P., & Barton, J. J. (2003). Bilateral deficits of transient visual attention in right parietal patients. *Brain, 126,* 2164–2174. [PubMed] [Article]

Becker, M. W., & Anstis, S. (2004). Metacontrast masking is specific to luminance polarity. *Vision Research, 44,* 2537–2543. [PubMed]

Blake, R., Tadin, D., Sobel, K. V., Raissian, T. A., & Chong, S. C. (2006). Strength of early visual adaptation depends on visual awareness. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 4783–4788. [PubMed] [Article]

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226,* 177–178. [PubMed]

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research, 13,* 767–782. [PubMed]

Cavanagh, P. (2004). Attention routines and the architecture of selection. In M. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 13–28). New York: Guilford Press.

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precueing of forced-choice discrimination. *Quarterly Journal of Experimental Psychology A, Human Experimental Psychology, 43,* 859–880. [PubMed]

Chung, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research, 41,* 1833–1850. [PubMed]

de Lange, H. (1958). Research into the dynamic nature of the human fovea–cortex systems with intermittent and modulated light: I. Attenuation characteristics with white and colored light. *Journal of the Optical Society of America, 48,* 777–784. [PubMed]

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143–149.

Eriksen, C. W., & Collins, J. F. (1969). Visual perceptual rate under two conditions of search. *Journal of Experimental Psychology, 80,* 489–492. [PubMed]

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383,* 334–337. [PubMed]

Hess, R. F., & Snowden, R. J. (1992). Temporal properties of human visual filters: Number, shapes and spatial covariation. *Vision Research, 32,* 47–59. [PubMed]

Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience, 4,* 127–128. [PubMed] [Article]

Huckauf, A., & Heller, D. (2004). On the relations between crowding and visual masking. *Perception & Psychophysics, 66,* 584–595. [PubMed]

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

Jacobs, R. J. (1979). Visual Resolution and contour interaction in the fovea and periphery. *Vision Research, 19,* 1187–1195. [PubMed]

Kahneman, D., & Henik, A. (1977). Effects of visual grouping on immediate recall and selective attention. In S. Dornic (Ed.), *Attention & performance VI* (pp. 307–332). Hillsdale, NJ: Erlbaum.

Kelly, D. H. (1961). Visual responses to time-dependent stimuli: I. Amplitude sensitivity measurements. *Journal of the Optical Society of America, 51,* 422–429. [PubMed]

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8,* 255–279. [PubMed]

Kremers, J., Kozyrev, V., Silveira, L. C., & Kilavik, B. E. (2004). Lateral interactions in the perception of flicker and in the physiology of the lateral geniculate nucleus. *Journal of Vision, 4*(7):10, 643–663, http://journalofvision.org/4/7/10/, doi:10.1167/4.7.10. [PubMed] [Article]

Lawrence, D. H. (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception & Psychophysics, 10,* 85–89.

Leat, S. J., Li, W., & Epp, K. (1999). Crowding in central and eccentric vision: The effects of contour interaction and attention. *Investigative Ophthalmology and Visual Science, 40,* 504–512. [PubMed] [Article]

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70,* 1458–1471. [PubMed]

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision, 2*(2):3, 167–177, http://journalofvision.org/2/2/3/, doi:10.1167/2.2.3. [PubMed] [Article]

Levi, D. M., Klein, S. A., & Hariharan, S. (2002). Suppressive and facilitatory spatial interactions in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision, 2*(2):2, 140–166, http://journalofvision.org/2/2/2/, doi:10.1167/2.2.2. [PubMed] [Article]

Mullen, K. T., & Losada, M. A. (1999). The spatial tuning of color and luminance peripheral vision measured with notch filtered noise masking. *Vision Research, 39,* 721–731. [PubMed]

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29,* 1631–1647. [PubMed]

Parkes, L., Lund, J., Angelucci, A., Solomon, J., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4,* 739–744. [PubMed]

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature detection and integration from detection. *Journal of Vision, 4*(12):12, 1136–1169, http://journalofvision.org/4/12/12/, doi:10.1167/4.12.12. [PubMed] [Article]

Poder, E. (2005). Effect of attention on the detection and identification of masked spatial patterns. *Perception, 34,* 305–318. [PubMed]

Potter, M. C. (1975). Meaning in visual search. *Science, 187,* 965–966. [PubMed]

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18,* 849–860. [PubMed]

Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review, 93,* 180–206. [PubMed]

Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1991). Phantom contours: Selective stimulation of the magnocellular pathways in man. *Investigative Ophthalmology and Visual Science, 32*(suppl. 26), 1034.

Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research, 38,* 71–77. [PubMed]

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 357–371. [PubMed]

Snowden, R. J., & Hess, R. F. (1992). Temporal frequency filters in the human peripheral visual field. *Vision Research, 32,* 61–72. [PubMed]

Spehar, B., & Zaidi, Q. (1997). Surround effects on the shape of the temporal contrast-sensitivity function. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 14,* 2517–2525. [PubMed]

Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review, 102,* 503–532.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11):8, 1024–1037, http://journalofvision.org/5/11/8/, doi:10.1167/5.11.8. [PubMed] [Article]

Thomas, J. P. (1985a). Detection and identification: How are they related? *Journal of the Optical Society of America A, 2,* 1457–1467. [PubMed]

Thomas, J. P. (1985b). Effect of static-noise and grating masks on detection and identification of grating targets. *Journal of the Optical Society of America A, 2,* 1586–1592. [PubMed]

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32,* 1349–1357. [PubMed]

Townsend, J. T., Taylor, S. G., & Brown, D. R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics, 10,* 375–378.

Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42,* 2357–2369. [PubMed]

van de Grind, W. A., Gruesser, O. J., & Lunkenheimer, H. U. (1973). Temporal transfer properties of the afferent visual system. Psychophysical, neurophysiological and theoretical investigations. In R. Jung (Ed.), *Handbook of sensory physiology, Vol. VII/3* (pp. 431–573). Berlin: Springer.

Verstraten, F. A., Cavanagh, P., & LaBianca, A. T. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research, 40,* 3651–3664. [PubMed]

Westheimer, G., Shimamura, K., & McKee, S. P. (1976). Interference with line-orientation sensitivity. *Journal of the Optical Society of America, 66,* 332–338. [PubMed]

Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 14,* 2057–2068. [PubMed]

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

Wolford, G., & Chambers, L. (1984). Contour interaction as a function of eccentricity. *Perception & Psychophysics, 36,* 457–460. [PubMed]