Crowding between first- and second-order letter stimuli in normal foveal and peripheral vision

Susana T. L. Chung

College of Optometry & Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX, USA

Roger W. Li

School of Optometry & Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA

Dennis M. Levi

School of Optometry & Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA

Evidence that the detection of first- and second-order visual stimuli is processed by separate pathways abounds. This study asked whether first- and second-order stimuli remain independent at the stage of processing where crowding occurs. We measured thresholds for identifying a first-order (luminance defined) or second-order (contrast defined) target letter in the presence of two second- or first-order flanking letters. For comparison, we also measured thresholds when the target and flanking letters were all first or second order. Contrast of the flankers was 1.6 times their respective contrast thresholds. Measurements were obtained at the fovea and 10° in the lower visual field of four normally sighted observers. Two observers were also tested at 10° nasal visual field. As expected, in both the fovea and periphery, the magnitude of crowding (threshold elevation) was maximal at the closest letter separation and decreased as letter separation increased. The magnitude of crowding was greater for second- than for first-order target letters, independent of the order type of flankers; however, the critical distance for crowding was similar for first- and second-order letters. Substantial crossover crowding occurred when the target and flanking letters were of different order type. Our finding of substantial interaction between first- and second-order stimuli suggests that the processing of these stimuli is not independent at the stage of processing at which crowding occurs.

Introduction

Our ability to see the details of a target is better when the target is presented alone than when the target is surrounded by other objects in proximity. This deleterious effect of proximal features on the visibility of the target, a form of inhibitory spatial interaction, is often referred to as crowding. Crowding is ubiquitous in spatial vision. For example, crowding affects two-bar resolution (Takahashi, 1967), Vernier discrimination (Levi & Klein, 1985; Levi, Klein, & Aitsebaomo, 1985; Westheimer & Hauske, 1975), stereopsis (Butler & Westheimer, 1978), line orientation sensitivity (Westheimer, Shimamura, & McKee, 1976), letter acuity (Flom, Weymouth, & Kahneman, 1963), and the contrast sensitivity for identifying letters (e.g., Chung, Levi, & Legge, 2001; Chung & Tjan, 2007; Pelli, Palomares, & Majaj, 2004; Strasburger, Harvey, & Rentschler, 1991).

Most previous work on crowding used targets that differed from their background by a change in luminance, often referred to as *luminance-defined* or *first-order* targets. In the absence of a change in luminance, targets can still

be distinguished from their background by a change in other stimulus attributes, such as contrast, texture, or motion. These stimuli are usually referred to as *second-order* stimuli. There is abundant psychophysical evidence that suggests that the spatial and temporal properties of first- and second-order visual perception are different. For instance, although both first- and second-order mechanisms are tuned for orientation and spatial frequency of the visual stimuli, the tuning bandwidth is wider for second- than for first-order stimuli (Landy & Oruc, 2002). Unlike that for first-order stimuli, modulation contrast sensitivity for second-order stimuli demonstrates weak dependence on the spatial frequency of the second-order modulating pattern (Landy & Oruc, 2002). Further, adaptation to first-order stimuli shows spatial-frequency and orientation selectivity, whereas adaptation to second-order stimuli only shows spatial-frequency selectivity but transfers across orientations (McGraw, Levi, & Whitaker, 1999). The inhibitory effect due to the presence of flanking Gabors on a target Gabor shows higher specificity with respect to spatial frequency and orientation for first- than for second-order stimuli (Ellemberg, Allen, & Hess, 2004). The spatial extent of interaction between these

Gabors extends over a larger distance for first- than for second-order stimuli (Ellemberg et al., 2004), although this finding seems unexpected based on the larger receptive field sizes for neurons processing second-order information (Mareschal & Baker, 1998; Rosa, 1997). With respect to temporal properties, visual evoked potential latencies are shorter and psychophysical reaction times are faster for first- than for second-order motion stimuli (Ellemberg et al., 2003). These results are in concordance with physiological (Baker & Mareschal, 2001; Mareschal & Baker, 1998) and brain-imaging (Dumoulin, Baker, Hess, & Evans, 2003; Larsson, Landy, & Heeger, 2006; Smith, Greenlee, Singh, Kraemer, & Hennig, 1998) evidence that there are two distinct processing streams for first- and second-order stimuli. Specifically, the processing of first-order information can be adequately represented by a linear filtering mechanism with a primary neural origin in V1. Processing of second-order information requires a non-linear stage, and the current model for such processing is a filter–rectifier–filter mechanism, with possible neural substrate in V2 (Mareschal & Baker, 1998; Schofield, 2000).

If the processing of first- and second-order objects were completely independent, then we would expect that observers' performance on psychophysical tasks would not be affected (improved or degraded) for stimuli that contain both first- and second-order information. Previous attempts to investigate spatial interactions between first- and second-order stimuli reveal conflicting findings. On one hand, several studies suggest no cross talk between first- and second-order information. For example, adaptation to either first- or second-order stimuli does not transfer to stimuli of the other type (Whitaker, McGraw, & Levi, 1997). Similarly, there is no crossover facilitation—a first-order grating pedestal does not facilitate the detection of second-order grating or vice versa (Schofield & Georgeson, 1999). Recently, we found that the performance for identifying second-order letters improves following perceptual learning of such letters, but the improvements do not transfer to the task of identifying first-order letters (Chung, Levi, & Li, 2006). More relevant to the present study, Banton and Levi (1993) found no crossover crowding between motion-defined (second order) targets and luminance-defined (first order) flanks (or vice versa) in a Vernier localization task.

On the other hand, a number of studies suggest substantial cross talk between first- and second-order information. For example, Rivest and Cavanagh (1996) showed substantial integration of different stimulus attributes (luminance, motion, color, and texture) in localizing contours. Using a simultaneous tilt illusion paradigm, Smith, Clifford, and Wenderoth (2001) showed that a vertical test grating appeared to tilt away from a surrounding, inducing grating even when the test and inducing gratings were of different order type (first or second order). For spatial-frequency discrimination, observers' performance for a grating stimulus that combined both first- and second-order information was better than predicted based on independent processing of the two types of stimulus

(Smith, & Scott-Samuel, 2001). Further, Ellemberg et al. (2004) found a substantial reduction in the apparent contrast of a second-order Gabor flanked by two first-order Gabors, although the effect for a first-order Gabor flanked by second-order Gabors was minimal.

A parsimonious explanation for the seemingly contradictory findings reviewed above is that the processing mechanisms and pathways for these two types of stimuli are separated at some stages (likely to be the early detection stage) but interact at some later stage, as suggested by Rivest and Cavanagh (1996). In this study, we asked whether first- and second-order stimuli remain independent at the stage of processing where crowding occurs. We also asked the question of whether crowding is stronger for second- than for first-order stimuli. To do so, we measured the magnitude of crowding for a first- or second-order target letter in the presence of two second- or first-order flanking letters, for various letter separations. For comparison, we also measured the magnitude of crowding when the target and flanking letters were all first or second order. To anticipate, we found that, in general, crowding is more substantial for second- than for first-order letters and that there is substantial crowding between first- and second-order letter stimuli, suggesting that the processing of first and second order information is not independent at the stage where crowding occurs.

Methods

To examine the question of whether first- and second-order letter stimuli remain independent at the stage of processing where crowding occurs, we compared the magnitude of crowding for identifying the middle target letters of trigrams (sequences of three lowercase letters) in which the target and flanking letters were combinations of first- or second-order letters, for various letter separations. The magnitude of crowding was quantified as the threshold elevation—threshold for identifying the target letters in the presence of flanking letters normalized to that for unflanked (single) letters. All four combinations of order types for target and flanking letters were examined (see Figure 1): a first-order target letter flanked by first-order letters (we shall refer to this condition as "111" for simplicity), a second-order target letter flanked by second-order letters ("222"), a first-order target letter flanked by second-order letters ("212"), and a second-order target letter flanked by first-order letters ("121"). We refer to the latter two as the crossover crowding conditions.

Stimuli

Our first-order letters are specified by a difference in luminance between the letter and its background, whereas
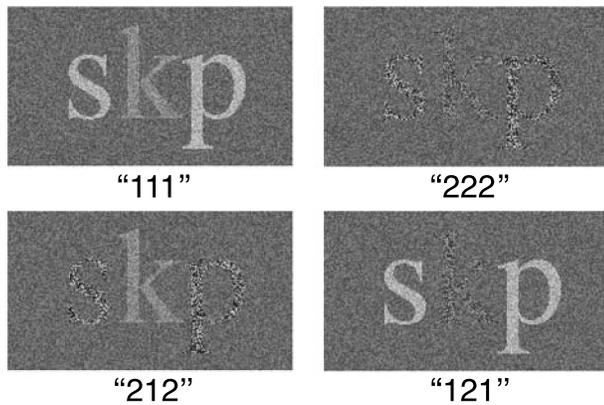
Figure 1. The trigram "skp" illustrated for the four testing conditions. The nickname by which we refer to each of the condition is given below each trigram: (top left) a first-order target (middle) letter flanked by two first-order letters, "111"; (top right) a second-order target letter flanked by second-order flankers, "222"; (bottom left) a first-order target letter flanked by second-order letters, "212"; (bottom right), a second-order target letter flanked by first-order letters, "121". Regardless of conditions, the task of the observer was to identify the middle letter of each trigram. During testing (not in this figure), the visibility of the flanking letters was matched between first- and second-order letters (see text for details).

second-order letters are specified by a difference in contrast between the letter and its background. Details for generating these first- and second-order letters are described elsewhere (Chung, Levi, et al., 2006). In brief, first-order (luminance-defined) letters were generated by assigning a different luminance value to the letter, compared with its midgray background. An array of white noise with an rms contrast of 0.07 covered both the letter and its background. Our primary measurement for first-order letters was the contrast threshold for identifying these letters, where contrast was defined as the Weber contrast between the letter and its background, (letter luminance − background luminance)/background luminance. Second-order (contrast-defined) letters were generated by assigning a different contrast to the white noise that made up the letter, with respect to the contrast of the background (rms contrast = 0.07). The mean luminance of the letter and that of its background were the same. Thus, threshold for identifying contrast-defined letters was defined as the just-noticeable differential contrast (ΔC) that defined the letter from its background, but for simplicity, we shall refer to this as the "contrast threshold" for second-order letters.

Psychophysical procedures

Each condition was tested in a separate block of trials. In each block of trials, we used the method of constant stimuli to present the letter that was to be identified (right or left

flanking letters when determining the thresholds for identifying the flanking letters or middle letters of the trigrams when determining the crowding effect) at five stimulus levels (five Weber contrast levels for first-order letters or five differential contrast [ΔC] for second-order letters), with each stimulus level presented 20 times within the block. Stimulus duration was 150 ms. Letters were randomly chosen with equal probability from the 26 lowercase letters of the Times Roman alphabet. Following the presentation of each trial, observers indicated their responses as to the identity of the letters using a keyboard. Audio feedback indicated whether or not the responses were correct. Data from each block of trials were fit with a cumulative Gaussian function from which threshold was defined as the stimulus level that yielded 50% correct identification of the letter (after correction for guessing). Data reported in this article represent the threshold averaged across four to six estimates (blocks) of the same condition. Testing was monocular with observers choosing their own preferred eye for foveal and 10° nasal field measurements. The non-tested eye was covered using a standard black eye patch. Testing was binocular for 10° lower field measurements.

Experimental design

To ensure that the letter size was not the limiting factor for observers' performance, we first measured the size thresholds for identifying single first-order letters of a fixed and high contrast of 0.7 and single second-order letters of a fixed ΔC of 0.7 at each of the tested retinal location (fovea, 10° lower field, and 10° nasal field). Letter size used in subsequent testing, regardless of whether the letters were first or second order, was set at 1.3 times the size threshold obtained for the second-order letters at a given retinal location and for all letter separations. On average, the size threshold for second-order letters was approximately six times larger than that for first-order letters, in both the fovea and periphery (Chung, Li, & Levi, 2006).

Given that second-order letters are less detectable than their first-order counterparts, to facilitate comparisons across conditions, thresholds for identifying the target letters were measured in the presence of flanking letters that were equated in visibility, regardless of whether they were first- or second-order letters. To do so, we first determined the contrast thresholds for identifying the flanking letters, for first- and second-order flankers, at each letter separation (see below). Thresholds for identifying the right and left flanking letters were determined in separate blocks of trials. On each trial, two letters that corresponded to the flanking letters of trigrams were presented (i.e., trigrams without the middle letters). Observers were instructed to maintain fixation at the fixation target (for peripheral testing) or in the middle of the two flanking letters if no fixation target was provided (for foveal testing). They were told before each block of trials which flanker

position (right or left) was the target. Except for the nasal visual field, thresholds for identifying the left and right flanking letters were very similar; thus, we used the value averaged across 8–12 blocks of trials (two letter positions and the 4–6 blocks for each letter position) to represent the identification threshold of the flanking letters. For the nasal visual field, thresholds for the right and left flanking letters were quite different; thus, we averaged their values separately across their respective 4–6 blocks of trials.

During trigram testing, the contrast (for first-order letters) or the differential contrast (for second-order letters) of the flanking letters was set at 1.6 times the respective threshold, as determined previously. We chose this value to equate for flanker visibility because this was the highest multiple of threshold that we could use to display the second-order letters due to the physical limitation of the range of contrast. At 1.6 times the threshold, the flanking letters were clearly visible to observers. Pelli et al. (2004) argued that once the flanker becomes visible (i.e., above detection threshold), its effect soon saturates, producing its full effect on the signal. Therefore, we were confident that the crowding effect would not have changed much had we increased (assuming that we were able to do so) the flanker visibility.

On each trial, three letters were presented at the center of the display, and the task of the observers was to identify the middle target letter. Three letter separations (defined as the center-to-center separation) were tested: $1\times$, $1.2\times$, and $2\times$ the height of the letter "x". At each letter separation, thresholds were determined for each of the four trigram conditions (111, 222, 212, and 121). For comparison, thresholds for identifying single first- and second-order letters were also measured. Note that at the smallest letter separation ($1\times$), occasionally, letter parts from adjacent letters (especially for "m" and "w") overlapped with one another. When this happened, the luminance value of each pixel represented the sum of the luminance values from the two letters, with clipping at the maximum contrast of 1.0. The deleterious effect of the overlapping was likely to be small, given that each threshold measurement for a given observer and a given trigram condition was based on hundreds of trials.

Apparatus

Stimuli were generated on a Macintosh G4 computer with software written in Matlab (The MathWorks, Massachusetts) using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and were displayed on a Sony 17-in. monitor (model number G400) at a mean luminance of 23 cd/m^2. The luminance of the display was measured using a Minolta photometer. By combining the red and blue output of the display with the use of a video attenuator (Pelli & Zhang, 1991) and custom-built software (Tjan, personal communication), we obtained an effective 10-bit resolution on luminance after correct-

ing for the gamma of the display. Observers were seated 42 cm from the display during testing (except when letter size thresholds were being determined). At this viewing distance, each pixel subtends 2.5 arcmin.

Observers

Four observers with normal vision, including two of the authors, participated in this study. Observers L.L. and S.Z. were unaware of the purpose of the study. All observers had best corrected visual acuity of 20/16 or better in both eyes and were experienced psychophysical observers. All observers, except for S.Z., had previous experience with our second-order letters (Chung, Levi, et al., 2006). Consequently, S.Z. was given an extensive period of practice (more than 10 hr) until his performance of identifying contrast-defined letters stabilized. Data collected during this practice phase are not included in this article. Written informed consent was obtained from each observer after the procedures of the experiment were explained and before the commencement of data collection.

Results

Fovea

For first-order stimuli (targets and flanks), crowding is absent or weak in the fovea (Chung et al., 2001; Levi, Klein, & Hariharan, 2002; Pelli et al., 2004), and this is evident in Figure 2. Here, we use threshold elevation to represent the magnitude of crowding. As expected for first-order letters, threshold elevation is maximal at the smallest letter separation and decreases with larger letter separation for the 111 condition. However, even at the smallest letter separation, crowding is small. Averaged across the four observers (±95% confidence intervals [CIs]), the magnitude of crowding is a mere 1.05 ± 0.06, consistent with the vast evidence that there is very little crowding in the normal fovea. In contrast, for second-order letters (the 222 condition), the magnitude of crowding is much larger (mean threshold elevation ± 95% CI = 1.34 ± 0.14) than that for the 111 condition at the smallest separation. With respect to our primary question of whether there is crowding when the target and flanking letters are of mixed-order type (the crossover conditions), our results show that in the fovea, crossover crowding is small (212: 1.06 ± 0.03; 121: 1.01 ± 0.06).

Besides the magnitude, crowding is also often characterized by its extent, generally represented by the spatial separation between the target and its flankers at which the magnitude of crowding drops to some criterion level. To determine if the spatial extent of crowding differs for the different trigram conditions, we fit the data for each trigram
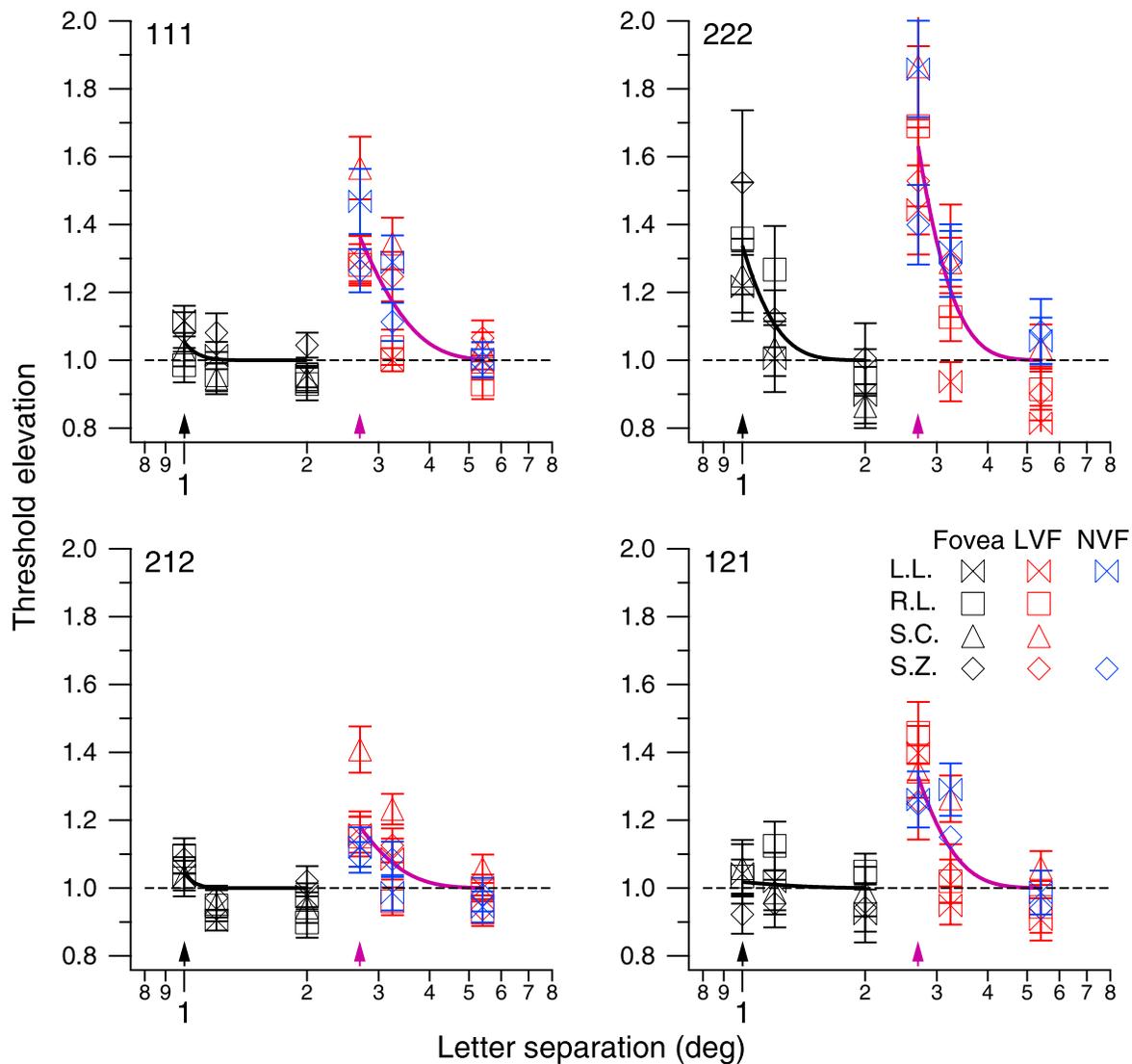
Figure 2. Threshold elevation—ratio of contrast threshold for identifying flanked and unflanked (single) letters—is plotted as a function of letter separation in degrees for the four trigram conditions. In each panel, data from individual observers are plotted as unfilled symbols (black: fovea; red: 10° lower field; blue: 10° nasal field). Solid lines represent the best fit curves (see text for details) through the foveal (black) or the combined peripheral (magenta) data. The black and magenta arrows represent the letter sizes (averaged across the observers) used at the fovea (1°) and at 10° eccentricity (both lower and nasal fields: 2.7°), respectively. At each retinal location (fovea or 10° eccentricity), three letter separations were tested, to which we refer as 1×, 1.2×, and 2× in the main text. Dashed lines represent the null effect (absence of threshold elevation). Error bars represent ±1 *SEM*.

condition with the following equation that describes a Gaussian function:

$$y = 1 + a \exp(1 - (x/b)^2), \qquad (1)$$

where a represents the peak amplitude of the function and b represents the x value at which peak amplitude occurs. This equation is a modification of the one used by Levi, Hariharan, and Klein (2002) and Levi, Klein, et al. (2002) in fitting their crowding data. In their studies, they used the "critical distance," defined as the distance between the target and its flanker that yielded a doubling in the

baseline (unflanked) threshold, to represent the spatial extent of crowding. Here, because we plotted threshold elevation on the y-axis, we defined the critical distance as the letter separation that corresponds to 2 *SD* from the peak of the Gaussian function. The critical distance was estimated based on the fitted parameters of the best fit curve, using the bootstrap resampling technique with 1,000 resamplings. Because the magnitude of crowding for the 111 and 121 conditions was not statistically significant from the null effect (no crowding), the critical distances for these conditions were not calculated. As summarized in Table 1, the critical distances for the 222 and 212 conditions are very similar and average 2.21°.

	Trigram condition			
	111	222	212	121
Fovea	No crowding	2.346 ± 0.005	2.064 ± 0.005	No crowding
10° Eccentricity	3.816 ± 0.003	3.430 ± 0.005	3.638 ± 0.005	3.494 ± 0.003

Table 1. Critical distances of crowding (±95% CI) as estimated based on the parameters of the best fit curves shown in Figure 2, using the bootstrapping resampling technique with 1,000 resamplings per estimate. Given that there is practically no crowding for the 111 and 121 conditions, the critical distances for these two conditions are omitted.

This value is relatively large for the crowding extent at the fovea, an issue that we will address in the Discussion section.

Periphery

As shown in Figure 2, there are no systematic or consistent differences between threshold elevations obtained at 10° lower (red) and nasal (blue) visual fields. Consequently, for the rest of the article, we will pool the data from the two retinal locations and simply refer to them as the peripheral data.

Consistent with previous studies for first-order letters that crowding is more substantial in the periphery than in the fovea (Chung et al., 2001; Levi, Hariharan, et al., 2002; Pelli et al., 2004), the magnitude of crowding for the 111 condition, at least at the smallest letter separation, is larger (mean threshold elevation ± 95% CI = 1.36 ± 0.10) than that obtained at the fovea. In fact, in the periphery, the averaged magnitude of crowding for each trigram con-

dition is larger than its foveal counterpart, at letter separations of 1× and 1.2×. It is also clear that crowding decreases with larger letter separations for all trigram conditions so that there is minimal crowding at a letter separation of 2×. Similar to the foveal data, the magnitude of crowding is also highest for the 222 condition (mean threshold elevation ± 95% CI = 1.63 ± 0.16). Unlike the foveal data, the magnitude of crowding for both crossover conditions is substantial, although there is an asymmetrical effect such that the magnitude is larger for the 121 than for the 212 condition (212: 1.18 ± 0.09; 121: 1.33 ± 0.07).

As expected, crowding also extends over a greater spatial extent in the periphery than in the fovea. There are no systematic or consistent differences in the critical distances for the four trigram conditions. Averaged across the four trigram conditions, the critical distance measures 3.59° at 10° eccentricity (see Table 1), a factor of 1.62× larger than the averaged extent at the fovea (2.21°).

To facilitate comparison of the magnitude of crowding across conditions and retinal locations, we plot the group-averaged threshold elevations (±95% CI) at the smallest
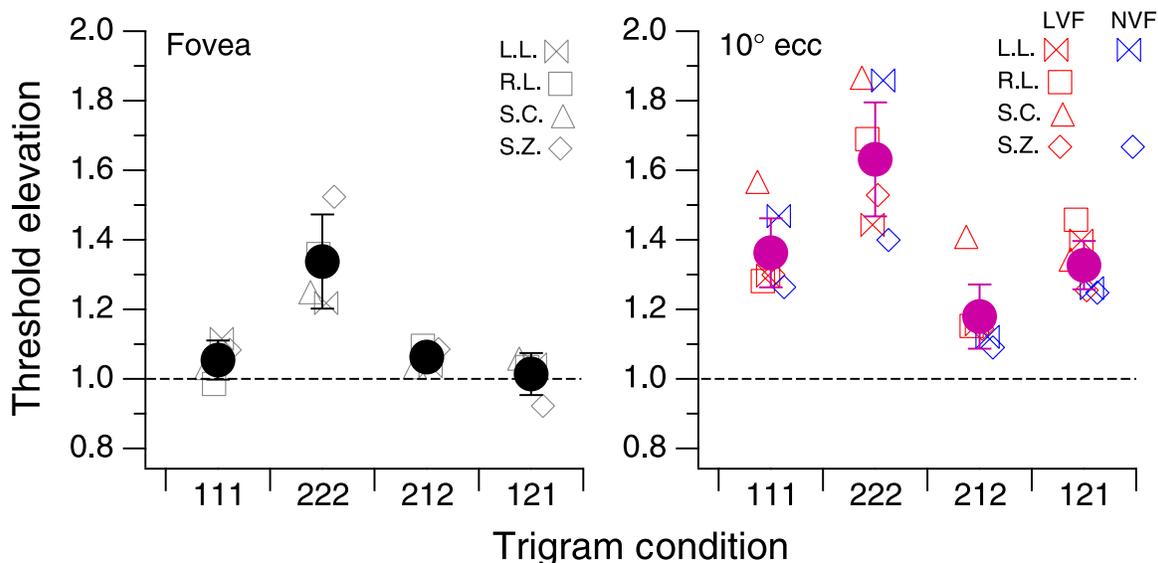

Figure 3. Crowding obtained at the smallest letter separation (1×) is compared for the four trigram conditions (111, 222, 212, and 121) at the fovea (left panel) and at 10° eccentricity (right panel: red, lower visual field [LVF]; blue, nasal visual field [NVF]). Individual observers' values are plotted as unfilled symbols of different shapes. Group-averaged values are represented by black (fovea) and magenta (periphery) filled circles, with the error bars representing the 95% CIs. Dashed lines represent the null effect (absence of threshold elevation).

letter separation (1×) in Figure 3 for the four trigram conditions, for the fovea and 10° eccentricity. A comparison of the values reveals three striking differences between data obtained at these two locations. First, and not surprisingly, crowding in the periphery is substantially larger than in the fovea for all trigram conditions. As pointed out earlier, this result simply reflects the more substantial crowding obtained in the periphery than in the fovea. Second, for all but the 212 condition, the magnitude of crowding in the periphery seems to be a constant shift from its respective value obtained at the fovea. This shift averages 0.3, implying that the magnitude of crowding obtained at 10° eccentricity is approximately 30% stronger than that obtained at the fovea, for the same trigram condition. Third, for the 212 condition, the magnitude of crowding in the fovea is slightly stronger than for the 111 and 121 conditions and is statistically different from a value of 1 (null effect), although the threshold elevation for both the 111 and the 121 conditions are not different from 1. In contrast, at 10° eccentricity, the magnitude of crowding for the 212 condition is weaker than for both the 111 and 121 conditions. Indeed, paired t tests comparing the 212 and the 111 or the 121 conditions show that pooled across observers, the threshold elevation at 10° eccentricity is lower for 212 than for 111, $t(6) = 5.53$, two-tailed $p = .003$, or 121, $t(6) = 2.87$, two-tailed $p = .035$.

Table 2 summarizes the threshold elevations (average ± 95% CI) for the various conditions at the smallest letter separation. An asterisk indicates that the value is statistically different from the null effect because the 95% CI includes the value of 1 (the null effect).

Discussion

The primary question we asked in this study was whether first- and second-order letter stimuli remain independent at the stage of processing where crowding occurs. We found substantial crossover crowding, suggesting that the processing of first- and second-order letters is not independent at the stage of processing where crowding takes place. There is an asymmetry in the magnitude of the crossover crowding: Crowding is stronger for a second-order target letter with first-order flankers (121) than for a first-order target letter with second-order flankers (212). The significance of this asymmetrical interaction will be discussed later. We also asked the question of whether crowding is stronger for second- than for first-order stimuli. Our finding of a much greater threshold elevation for second-order letters (when the target and its flanking letters were all second order) than for first-order letters implies that the strength of crowding is much stronger for second- than for first-order stimuli. However, there is no difference in the critical distances for crowding between first- and second-order letters. In the following subsections, we will discuss our findings in relation to previous studies and present our view of how our data may contribute to the overall understanding of the processing of first- and second-order information.

Crowding of second-order letters

By comparing the threshold for identifying a target letter in the presence of first- or second-order letters that were equated in visibility (1.6 times their respective identification threshold), we found that the threshold elevation is greater for second- than for first-order letters. We interpreted this finding as an indication that crowding is more substantial for second- than for first-order letters. However, if the contrast response functions for identifying first- and second-order letters have different slopes, then simply setting the flanker contrast as 1.6 times their respective identification thresholds would have produced first- and second-order flankers that were not matched in visibility. Wong, Levi, and McGraw (2001) measured detection thresholds for first- and second-order Gabor stimuli by presenting the stimuli at various contrast levels. They found that for observers with normal vision, the averaged slope of the psychometric functions for detecting second-order targets was shallower than that for detecting first-order targets by about a factor of 1.13. Landy and Oruc (2002) reported that the modulation depth increment threshold functions are much shallower for second- than for first-order gratings. In our study, we also found that the slope of the psychometric functions relating percent-correct and contrast was shallower for identifying second- than first-order letters by approximately a factor of 1.27, independent of whether measurements were obtained at the fovea or at 10° eccentricity. In other words, it is likely that using the same fixed multiple of threshold may not have equated the flanker visibility perfectly. How would

| | Trigram condition | | | |
	111	222	212	121
Fovea	1.054 ± 0.056	1.338 ± 0.135*	1.063 ± 0.032*	1.014 ± 0.061
10° Eccentricity	1.363 ± 0.100*	1.631 ± 0.164*	1.179 ± 0.092*	1.327 ± 0.070*

Table 2. Averaged values (±95% CI) of the magnitude of crowding for the different testing conditions shown in Figure 3. Asterisks indicate values that are statistically different from the null effect (no crowding) at $p < .05$.

this affect the interpretation of our findings? First, Pelli et al. (2004) showed that as soon as the flankers become visible, the effect of crowding soon saturates; therefore, the slight difference in visibility between our first- and second-order letters, even if they had not been equated properly, might not have produced sizeable difference in the crowding effect. Second, by setting the flanker contrast at 1.6 times their respective identification threshold, the percent-correct for identifying second-order flankers was *lower* than that for identifying first-order flankers because of the shallower slope. Given that the magnitude of crowding increases with flanker contrast (Chung et al., 2001; Pelli et al., 2004), the crowding effect is expected to be *lower* for second- than for first-order flankers. However, we found the opposite result (222 condition vs. 121 condition, at both the fovea and 10° eccentricity), implying that, if anything, we might have underestimated how much stronger the magnitude of crowding is for second- than for first-order letters.

Schofield and Georgeson (1999) showed that although second-order signals demonstrate facilitation for the task of detection, they do not mask one another. Based on these data, one might predict that there would be no crowding for second-order letters. Contrary to this prediction, we found substantial crowding for second-order letters. The difference between our study and that of Schofield and Georgeson could be due to the different tasks involved. Schofield and Georgeson used a contrast detection task whereas ours was a suprathreshold letter identification task. Indeed, by measuring the apparent contrast reduction of a Gabor target in the presence of two flanking Gabors, also a suprathreshold task, Ellemberg et al. (2004) showed substantial interaction among second-order targets.

Although the magnitude of crowding is larger for second- than for first-order letters, the critical distances for crowding are very similar between the two types of letters. In fact, at a given retinal location, the critical distance for crowding seems to be a constant regardless of the trigram condition—2.21° and 3.59° at the fovea and at 10° eccentricity, respectively. How these values compare with those in the literature and the significance of their meaning will be discussed in a later section. For now, the important point is that the critical distance for crowding does not depend on the order type (first or second order) of the target or flanking letters. Given that crowding is believed to result from a combination or pooling of signals from the target and its flankers after the stage of feature extraction (Levi, Hariharan, et al., 2002; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli et al., 2004), our result implies that the receptive field properties at the combination site do not care about the order type of letters. Note that although we found that the critical distance for crowding does not vary with the order type of letters, Ellemberg et al. (2004) reported a smaller extent of interaction for their second-order targets for the task of apparent contrast perception. It is unclear why they found such an effect, unless it is due solely to the specific task they used.

Differences between the fovea and periphery

Consistent with previous studies, we found that both the magnitude and the extent of crowding are larger in the periphery than in the fovea (Figure 2). However, previous studies have only examined crowding for first-order targets. Here, we extended the same finding qualitatively to second-order letters and in the presence of mixed first- and second-order letters. We also found that the magnitude of crowding seems to be a constant amount (approximately 30%) higher at 10° eccentricity (both lower and nasal fields) than at the fovea, independent of trigram conditions (except for the 212 condition, which we shall address in the following paragraph). The larger magnitude of crowding found in the periphery than at the fovea, for both first- and second-order letters, could be related to the larger critical distance for crowding (larger receptive fields or greater pooling of neural signals) corresponding to the peripheral rather than the foveal locations. It may also reflect the fact that the stimuli were larger in the periphery than in the fovea (by approximately 2.7×).

An interesting result was that for the 212 condition (first-order letter flanked by second-order letters), the magnitude of crowding was slightly *higher* than those for its 111 and 121 counterparts at the fovea but *lower* than those for the 111 and 121 conditions in the periphery. This is simply explained if the flanking effect is always stronger when the flankers are of the same order type as that of the target, and weaker when the flankers are of a different order type (crossover crowding). This explanation is consistent with the results obtained for the second-order target letters—the magnitude of crowding is smaller for the 121 than for the 222 condition, in both the fovea and periphery. For first-order target letters, the result obtained in the periphery is also consistent with the explanation (magnitude of crowding is smaller for the 212 than for the 111 condition). At the fovea, however, there is practically no crowding for the 111 condition, leaving almost no room for the 212 condition to have an even smaller threshold elevation. It is noteworthy that theoretically, we could obtain a magnitude of crowding <1, implying facilitation due to the presence of flankers. However, we have previously shown that unlike contrast detection, letter crowding does not demonstrate facilitation (Chung et al., 2001).

We found that the critical distances for crowding are 2.21° and 3.59° at the fovea and at 10° eccentricity, respectively. The value at the fovea is much larger than that reported previously (e.g., 0.5°: Chung et al., 2001; 0.06°: Toet & Levi, 1992) and is likely due to the large letter size we used (1° letters) because the critical distance for crowding in the fovea is proportional to target size (Levi, Klein, et al., 2002). Alternatively, the critical distances estimated at the fovea were derived from only two trigram conditions, both of which involved second-order letters (212 and 222). Therefore, it is likely that the

larger-than-expected critical distance at the fovea could be simply explained as a larger integration or pooling zone for second-order letters. However, the highly similar critical distances obtained for the four trigram conditions at 10° eccentricity suggest that the integration zone does not depend on the order type of letters (see the next section). At 10° eccentricity, the average critical distance is approximately 0.4 times the eccentricity, very close to the value reported by Bouma (1970), and is in line with the values reported by others (see Chung et al., 2001, for a review).

Do first- and second-order information interact?

As we briefly summarized in the Introduction, there is conflicting psychophysical evidence as to whether the processing of first- and second-order information remains completely independent. Clearly, there are substantial methodological differences among these studies that may account for the contradictory results of the studies. Here, we will discuss some of these differences.

Allen, Hess, Mansouri, and Dakin (2003) suggested that when a stimulus is defined by both first- and second-order signals, for example, a first-order carrier windowed by a second-order envelope such as a Gabor patch, the two types of information can hardly be considered as separate because both types of information are required to represent the stimulus. As such, findings of an interaction between first- and second-order information from studies that employed this type of stimulus (e.g., Dakin, Williams, & Hess, 1999; Morgan, Mason, & Baldassi, 2000) cannot be used as evidence refuting the independent processing of first- and second-order information. However, this explanation cannot readily explain the findings of other studies in which combinations of first- and second-order stimuli are spatially separated, as in the present study.

An examination of previous studies reveals that the contradictory results among different studies regarding the interaction between first- and second-order spatial targets could be, in part, attributed to the task. Specifically, studies that involve a detection task generally yield results showing that the detection of first- and second-order stimuli is completely independent and that they do not interact with one another (Schofield & Georgeson, 1999). In contrast, tasks that involve suprathreshold judgments such as spatial-frequency discrimination (Smith, & Scott-Samuel, 2001), perceived position judgment (Rivest & Cavanagh, 1996; Whitaker, McGraw, Keeble, & Skillen, 2004), apparent contrast perception (Ellemberg et al., 2004), and tilt illusion (Smith, Clifford, et al., 2001) tend to show interaction between first- and second-order signals. In the present study, we used a letter identification task and also found substantial interaction between first- and second-order letters. Our finding is augmented by the study of Oruc, Landy, and Pelli (2006) who used a noise-

masking paradigm to derive the "channel" for identifying texture-defined letters (also a second-order task). They found that the identification of second-order letters shares several important properties of that for identifying first-order letters, but one difference is that second-order letter identification is scale invariant whereas first-order letter identification is not. Their results imply that identification of first- and second-order letters is not entirely independent (but not identical either).

However, task per se apparently cannot fully predict whether or not interaction exists between first- and second-order targets. Using a position adaptation paradigm, Whitaker et al. (1997) found minimal shift in the perceived position of a target (first or second order) that was presented subsequent to an adapting target of the other order type. One possible explanation for their failure to find an interaction effect despite their use of a suprathreshold task is that adaptation to their first-order targets, luminance-defined Gaussian blobs, might have occurred at the retinal instead of at the cortical level. If so, then it is not surprising that adapting to their first-order Gaussian blobs did not affect the subsequently presented second-order texture-defined patches that are, presumably, processed in extrastriate cortical areas and vice versa. Another exception is the finding that there is no crossover crowding between motion-defined (second order) and luminance-defined (first order) stimuli in a Vernier task (Banton & Levi, 1993). A plausible explanation for the findings of Banton and Levi (1993) and Whitaker et al. (1997) is that the coding of perceived location and the precision of the coding may be performed by different mechanisms (Rivest & Cavanagh, 1996).

The most parsimonious account for the seemingly conflicting psychophysical evidence is that first- and second-order processing remains separate at the detection level, but the signals are combined at a later processing stage (Rivest & Cavanagh, 1996). This explanation can account for much of the previous data on cross talk. It also provides an explanation for the lack of a transfer of learning from second-order to first-order targets in our previous study (Chung, Levi, et al., 2006) because the first- and second-order signals were never presented simultaneously and were, therefore, never combined.

Where are these signals combined? Physiological studies point to cortical area V4 as a plausible site for the combination of signals from different stimulus types (Ferrera, Nealey, & Maunsell, 1992, 1994; Logothetis & Charles, 1990), which is consistent with our findings that the neurons at the site where crowding occurs do not care about whether the letters are first or second order. More relevant to our study, V4 has been implicated for crowding in macaque monkeys (Motter, 2002). Smith, Singh, Williams, and Greenlee (2001) reported that the receptive field size in V4 increases from approximately 2.5° at the fovea to 7° at 10° eccentricity. Our psychophysical estimate of the critical distance for crowding averages 2.21° at the fovea and 3.59° at 10° eccentricity.

However, our critical distance refers to the separation between the target letter and one of its flanking letters. If we assume that the spatial extent of crowding is symmetrical on either side of the target letter, then the total extent can be approximated by 2× our critical distance. Figure 4 plots such extent for the four trigram conditions obtained at the fovea and at 10° eccentricity. The dashed lines (black for fovea and magenta for 10° eccentricity) represent the receptive field size in V4 as reported by Smith, Singh et al. (2001, see their Figure 9). Clearly, our empirical data for the extent of crowding match very closely with the receptive field size of V4 neurons at 10° eccentricity but not at the fovea. This difference could be due to the different mechanisms underlying foveal and peripheral crowding. Levi, Hariharan, et al. (2002) and Levi, Klein, et al. (2002) suggested that foveal crowding is simply contrast masking, whereas peripheral crowding represents the genuine crowding effect and is limited by the pooling of signals from the target and flankers at a second stage, after an initial stage of feature extraction. As such, our estimates of the critical distances in the fovea and periphery may not be directly comparable because they reflect the resolution limit of different mechanisms.

Asymmetrical crossover effects

We found significant crossover of crowding for first-order letters flanked by second-order letters (the 212 condition) and vice versa (the 121 condition), at least in the periphery where crowding is strong. However, the magnitude of effect is not the same for the two conditions.
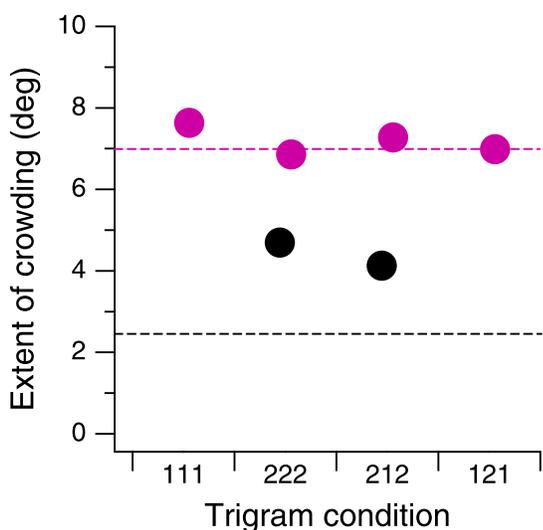

Figure 4. The spatial extent of crowding (2× the critical distance, in degrees) is plotted as a function of trigram condition for the fovea (black circles) and 10° eccentricity (magenta circles). Dashed lines represent the receptive field sizes in V4 as reported by Smith, Singh, et al. (2001) at these two retinal locations (black: fovea; magenta: 10° eccentricity).

Specifically, the magnitude of crowding is larger for second-order letters flanked by first-order letters (121) than for first-order letters flanked by second-order letters (212). A similar asymmetrical crossover interaction effect has been reported previously for other psychophysical tasks. Ellemberg et al. (2004) found that spatially separated first-order Gabor flankers cause a larger reduction in the apparent contrast of a flanked second-order target Gabor, compared with the effect caused by second-order flanking Gabor on a first-order target Gabor. For the tilt illusion, Smith, Clifford, et al. (2001) found that the repulsion effect (an inducer grating causing the test grating to appear to tilt away from its veridical orientation) is larger for a second-order test grating surrounded by a first-order inducer, compared with a first-order test grating surrounded by a second-order inducer.

Given that the asymmetrical crossover effect is found for other spatial tasks, it may represent a generic property of the interaction between the two pathways processing first- and second-order visual information. The stronger interaction effect found when the target in question is a second-order target with the interaction coming from first-order objects, compared with the reverse case, could be simply due to a different weighting of the first- versus second-order signals. In the following section, we will present a framework for how first- and second-order objects may interact.

How do first- and second-order information interact?

To account for the asymmetrical crossover interaction effects, as well as the differences between the foveal and peripheral results, we suggested in previous sections that the visual system assigns different weighting to different targets. Here, we provide a framework for the combination of first- and second-order stimuli and compare the predictions generated from this framework with our empirical results.

We begin with the assumption that the visual system generally places more weighting on first- than on second-order signals, perhaps because first-order signals are usually more salient. We further assume that signals from targets are weighted more than signals from flankers. Thirdly, we assume that at the locus of crowding, the combination rule gives flankers that are of the same order type as the target more weight than flankers of a different order type. These assumptions lead to several predictions. First, because first-order signals are generally weighted more than second-order signals, it follows that, in general, first-order target letters suffer from less crowding than second-order target letters. Our results (Figures 2 and 3) are consistent with this prediction. Second, for first-order targets, the flankers would be more effective (because the signals are weighted more) for first-order (same order) than for second-order (different) flankers; thus, the

prediction is that there is more crowding for the 111 condition than for the 212 condition. Our results obtained in the periphery are consistent with this prediction. As for the foveal data, as noted in the previous section, the reason we did not observe *less* crowding for the 212 condition as compared with the 111 condition is likely due to a floor effect. Similarly, for second-order targets, second-order flankers would be more effective than first-order flankers, predicting that the 222 condition would yield more crowding than the 121 condition. Our results are consistent with this prediction at both the fovea and 10° eccentricity. In fact, at the fovea, the magnitude of crowding for the 222 condition was the only one that was clearly different from the null effect (Figure 3).

It is worth nothing that our assumptions above predict that the 222 trigram condition should yield the strongest crowding among the four trigram conditions. Figure 3 shows that this is indeed the case, at both the fovea and 10° eccentricity. As pointed out above, there is very little crowding at the fovea in general, yet the 222 condition still shows quite substantial crowding—an observation consistent with the first and third assumptions of our framework.

Clearly, the assumptions of the framework presented above need to be fleshed out and validated quantitatively to provide a more complete picture of how first- and second-order targets interact. Nevertheless, in this study, we have clearly demonstrated that second-order signals can inhibit one another and that the processing of first- and second-order letters is not independent, at least not at the stage of processing at which crowding occurs.

Acknowledgments

Corresponding author: Susana T. L. Chung.
Email: schung@optometry.uh.edu.
Address: College of Optometry, University of Houston, Houston, TX 77204-2020, USA.

References

Allen, H. A., Hess, R. F., Mansouri, B., & Dakin, S. C. (2003). Integration of first- and second-order orientation. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 20,* 974–986. [PubMed]

Baker, C. L., Jr., & Mareschal, I. (2001). Processing of second-order stimuli in the visual cortex. *Progress in Brain Research, 134,* 171–191. [PubMed]

Banton, T., & Levi, D. M. (1993). Spatial localization of motion-defined and luminance-defined contours. *Vision Research, 33,* 2225–2237. [PubMed]

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

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Butler, T. W., & Westheimer, G. (1978). Interference with stereoscopic acuity: Spatial, temporal, and disparity tuning. *Vision Research, 18,* 1387–1392. [PubMed]

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

Chung, S. T. L., Levi, D. M., & Li, R. W. (2006). Learning to identify contrast-defined letters in peripheral vision. *Vision Research, 46,* 1038–1047. [PubMed]

Chung, S. T. L., Li, R. W., & Levi, D. M. (2006). Identification of contrast-defined letters benefits from perceptual learning in adults with amblyopia. *Vision Research, 46,* 3853–3861. [PubMed]

Chung, S. T. L., & Tjan, B. S. (2007). Shift in spatial scale in identifying crowded letters. *Vision Research, 47,* 437–451. [PubMed]

Dakin, S. C., Williams, C. B., & Hess, R. F. (1999). The interaction of first- and second-order cues to orientation. *Vision Research, 39,* 2867–2884. [PubMed]

Dumoulin, S. O., Baker, C. L., Jr., Hess, R. F., & Evans, A. C. (2003). Cortical specialization for processing first- and second-order motion. *Cerebral Cortex, 13,* 1375–1385. [PubMed][Article]

Ellemberg, D., Allen, H. A., & Hess, R. F. (2004). Investigating local network interactions underlying first- and second-order processing. *Vision Research, 44,* 1787–1797. [PubMed]

Ellemberg, D., Lavoie, K., Lewis, T. L., Maurer, D., Lepore, F., & Guillemot, J. P. (2003). Longer VEP latencies and slower reaction times to the onset of second-order motion than to the onset of first-order motion. *Vision Research, 43,* 651–658. [PubMed]

Ferrera, V. P., Nealey, T. A., & Maunsell, J. H. (1992). Mixed parvocellular and magnocellular geniculate signals in visual area V4. *Nature, 358,* 756–761. [PubMed]

Ferrera, V. P., Nealey, T. A., & Maunsell, J. H. (1994). Responses in macaque visual area V4 following inactivation of the parvocellular and magnocellular LGN pathways. *Journal of Neuroscience, 14,* 2080–2088. [PubMed] [Article]

Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual resolution and contour interaction. *Journal of the Optical Society of America, 53,* 1026–1032. [PubMed]

Landy, M. S., & Oruc, I. (2002). Properties of second-order spatial frequency channels. *Vision Research, 42,* 2311–2329. [PubMed]

Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology, 95,* 862–881. [PubMed] [Article]

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. (1985). Vernier acuity, crowding and amblyopia. *Vision Research, 25,* 979–991. [PubMed]

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research, 25,* 963–977. [PubMed]

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]

Logothetis, N. K., & Charles, E. R. (1990). The minimum motion technique applied to determine isoluminance in psychophysical experiments with monkeys. *Vision Research, 30,* 829–838. [PubMed]

Mareschal, I., & Baker, C. L., Jr. (1998). A cortical locus for the processing of contrast-defined contours. *Nature Neuroscience, 1,* 150–154. [PubMed] [Article]

McGraw, P. V., Levi, D. M., & Whitaker, D. (1999). Spatial characteristics of the second-order visual pathway revealed by positional adaptation. *Nature Neuroscience, 2,* 479–484. [PubMed] [Article]

Morgan, M. J., Mason, A. J., & Baldassi, S. (2000). Are there separate first-order and second-order mechanisms for orientation discrimination? *Vision Research, 40,* 1751–1763. [PubMed]

Motter, B. C. (2002). Crowding and object integration within the receptive field of V4 neurons [Abstract]. *Journal of Vision, 2*(7):274, 274a, http://journalofvision.org/2/7/274/, doi:10.1167/2.7.274.

Oruc, I., Landy, M. S., & Pelli, D. G. (2006). Noise masking reveals channels for second-order letters. *Vision Research, 46,* 1493–1506. [PubMed]

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. [PubMed]

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

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research, 31,* 1337–1350. [PubMed]

Rivest, J., & Cavanagh, P. (1996). Localizing contours defined by more than one attribute. *Vision Research, 36,* 53–66. [PubMed]

Rosa, M. (1997). Visuotopic organization of primate extrastriate cortex. In K. S. Rockland, J. H. Kaas, & A. Peters (Eds.), *Cerebral cortex: Extrastriate cortex in primates* (vol. 12, pp. 127–203). New York: Plenum Press.

Schofield, A. J. (2000). What does second-order vision see in an image? *Perception, 29,* 1071–1086. [PubMed]

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behaviour. *Vision Research, 39,* 2697–2716. [PubMed]

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M. & Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *Journal of Neuroscience, 18,* 3816–3830. [PubMed] [Article]

Smith, A. T., & Scott-Samuel, N. E. (2001). First-order and second-order signals combine to improve perceptual accuracy. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 18,* 2267–2272. [PubMed]

Smith, A. T., Singh, K. D., Williams, A. L., & Greenlee, M. W. (2001). Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cerebral Cortex, 11,* 1182–1190. [PubMed] [Article]

Smith, S., Clifford, C. W., & Wenderoth, P. (2001). Interaction between first- and second-order orientation channels revealed by the tilt illusion: Psychophysics and computational modelling. *Vision Research, 41,* 1057–1071. [PubMed]

Strasburger, H., Harvey, L. O., Jr., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics, 49,* 495–508. [PubMed]

Takahashi, E. S. (1967). *Effects of flanking contours on visual resolution at foveal and near-foveal loci.*

Unpublished doctoral dissertation, University of California, Berkeley.

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

Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research, 15,* 1137–1141. [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]

Whitaker, D., McGraw, P. V., Keeble, D. R., & Skillen, J. (2004). Pulling the other one: 1st- and 2nd-order visual information interact to determine perceived location. *Vision Research, 44,* 279–286. [PubMed]

Whitaker, D., McGraw, P. V., & Levi, D. M. (1997). The influence of adaptation on perceived visual location. *Vision Research, 37,* 2207–2216. [PubMed]

Wong, E. H., Levi, D. M., & McGraw, P. V. (2001). Is second-order spatial loss in amblyopia explained by the loss of first-order spatial input? *Vision Research, 41,* 2951–2960. [PubMed]