

The effect of crowding on orientation-selective adaptation in human early visual cortex

Taiyong Bi

Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, P.R. China



Peng Cai

State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, P.R. China, & Graduate University, Chinese Academy of Sciences, Beijing, P.R. China



Tiangang Zhou

State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, P.R. China



Fang Fang

Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, P.R. China



Crowding is the identification difficulty for a target in the presence of nearby flankers. Based on psychophysical findings, many theories have been proposed to explain crowding at multiple levels. However, little is known about its neural mechanism. In this study, we combined psychophysical and fMRI adaptation techniques to search for the cortical locus of crowding. In the psychophysical experiment, when subjects' attention was controlled, we found that the threshold elevation aftereffect (TEAE) was not affected by crowding, regardless of the contrast level of adapting stimulus. In the fMRI experiment, the orientation-selective fMRI adaptation in V1 was not affected by crowding either. However, downstream from V1, we found that crowding weakened the adaptation effect in V2 and V3. Our results demonstrate that crowding occurs beyond V1 and provide one of the first pieces of direct evidence supporting the two-stage model of crowding (D. M. Levi, 2008).

Keywords: adaptation, contrast, crowding, fMRI, orientation, visual cortex

Citation: Bi, T., Cai, P., Zhou, T., & Fang, F. (2009). The effect of crowding on orientation-selective adaptation in human early visual cortex. *Journal of Vision*, 9(11):13, 1–10, <http://journalofvision.org/9/11/13/>, doi:10.1167/9.11.13.

Introduction

When a peripheral target is presented with nearby flankers, it is much harder to identify. This phenomenon is referred to as crowding. Crowding is prevalent in spatial vision and occurs under a wide range of conditions and tasks, including Vernier acuity (Levi, Klein, & Aitsebaomo, 1985; Williams & Esworth, 1986), stereoacuity (Westheimer & Truong, 1988), orientation discrimination (Westheimer, Shimamura, & McKee, 1976), contrast discrimination (Wilkinson, Wilson, & Ellernberg, 1997), letter recognition (Nandy & Tjan, 2007; Pelli, Palomares, & Majaj, 2004), and face recognition (Louie, Bressler, & Whitney, 2007; Martelli, Majaj, & Pelli, 2005). The importance of studying crowding in vision sciences has recently been recognized more because it may help us understand the processes (e.g., feature integration) critical to object recognition (Levi, 2008).

Although crowding has been extensively studied using visual psychophysics for more than 80 years (Bouma, 1970; Korte, 1923; Stuart & Burian, 1962), little is known about its neural mechanism. Based on psychophysical findings, various theories and ideas have been proposed to explain crowding at multiple levels. The “optics” explanation attributes crowding to the effect of the eye’s point spread function when stimuli are small and closely spaced (Hess, Dakin, & Kapoor, 2000; Liu & Arditi, 2000). The receptive field theory suggests that crowding occurs when a target and flankers fall within a single receptive field and reflects pooling of them by the receptive field (Flom, Heath, & Takahashi, 1963). A related idea is that crowding is due to long-range horizontal connections between neurons with similar tuning properties in early visual cortex (e.g., V1) (Gilbert, 1998). On the other hand, attention theories argue that crowding could be ascribed to coarse resolution of spatial attention (He, Cavanagh, & Intriligator, 1996, 1997; Intriligator & Cavanagh, 2001) or

unfocussed spatial attention (Strasburger, Harvey, & Rentschler, 1991; Strasburger, 2005). None of these theories could provide an adequate explanation for the large body of existing psychophysical phenomena (see a comprehensive review by Levi, 2008).

A straightforward way to reveal the neural mechanism of crowding is searching for the neural locus of crowding. Surprisingly, to the best of our knowledge, only two fMRI studies, including one published abstract, have been carried out to investigate this issue directly. Arman, Chung, and Tjan (2006) found that crowding could modulate BOLD signal as early as in V2, but not in V1. Fang and He (2008) demonstrated that V1 BOLD signal to the target was not affected by the presence of flankers; instead, the distribution of attentional modulation was influenced by the crowding effect. Thus, both studies suggest that crowding occurs beyond V1, which is consistent with previous psychophysical findings (Chung, Li, & Levi, 2007; He et al., 1996; Liu, Jiang, Sun, & He, 2009; Tripathy & Cavanagh, 2002). On the other hand, Blake, Tadin, Sobel, Raissian, and Chong (2006) showed that the threshold-elevation aftereffect (TEAE) from adapting to a low-contrast grating (not a high-contrast grating) was significantly reduced by crowding and argued that the full-strength TEAE during crowding reported by He et al. (1996) could be explained by the saturation of aftereffect strength at a high adapting contrast level. Their results imply that crowding starts at an early stage of cortical processing (e.g., V1) because the TEAE originates, at least in part, from V1 adaptation.

When we utilize Blake and colleagues' finding to make inferences about the cortical locus of crowding, two issues need to be considered carefully. First, the reduced TEAE at the low adapting contrast level might be attributed to less spatial attention attracted to the target because the presence of flankers reduced its visibility. It has been demonstrated that spatial attention could enhance contrast responses in human visual cortex (Buracas & Boynton, 2007; Li, Lu, Tjan, Doshier, & Chu, 2008) and early visual adaptation (Festman & Ahissar, 2004; Spivey & Spirn, 2000). Second, psychophysical data alone do not allow us to determine in which visual cortical area crowding is inaugurated because contrast and orientation adaptations are evident from V1 to V4 (Fang, Murray, Kersten, & He, 2005; Gardner et al., 2005). V1, V2, and V4 have been proposed to be the cortical locus of crowding (Arman et al., 2006; Blake et al., 2006; Levi, 2008); however, none of these proposals has been substantiated clearly by neuroimaging studies in human subjects.

To address these issues, we first measured the TEAE both with and without crowding. Then we used high-resolution fMRI to quantify orientation-selective adaptation in early visual cortical areas and examined how crowding influenced the adaptation effect in different areas. In the experiments, subjects' attention was controlled by asking

them to do a fixation task (Gardner et al., 2005). Both high- and low-contrast gratings were used as adapting stimuli.

Methods

Participants

A total of five healthy subjects (four male and one female) participated in all the experiments. They were right-handed with reported normal or corrected-to-normal vision and had no known neurological or visual disorders. Their ages ranged from 21 to 34 years. They gave written, informed consent in accordance with the procedures and protocols approved by the human subjects review committee of Peking University.

Stimuli and design

Four adapting stimuli were used in both psychophysical and fMRI experiments (Figure 1A). In the unflanked condition, only a circular patch of a sinusoidal grating (radius: 1.5° ; spatial frequency: 2 cycles/ $^\circ$; contrast: 0.9 or 0.14; orientation: $\pm 45^\circ$) was presented as a target, which was centered at 7° eccentricity in the left visual field. In the flanked condition, the grating was immediately flanked by two circular patches with radii of 1° and 2.2° . The flankers were made of two overlaid orthogonal sinusoidal gratings (spatial frequency: 2 cycles/ $^\circ$; contrast: 0.45; orientation: $\pm 45^\circ$) and were centered at 4.5° and 10.7° eccentricities, respectively. The areas of the target and flankers were scaled for cortical magnification (Duncan & Boynton, 2003). The adapting stimuli flickered in counter-phase at 2 Hz in the experiments.

For the psychophysical experiments, stimuli were displayed on a uniform gray field (48 cd/m^2 , mean luminance of the target grating) at a viewing distance of 57 cm. The display was a IIYAMA color graphic monitor (model: HM204DT; refresh rate: 85 Hz; resolution: 1024×768 ; size: 22 inches). The stimuli were rendered with a video card with 8-bit input resolution and 14-bit output resolution using Cambridge Research System Bits++. The output luminance of the monitor was linearized using a look-up table in conjunction with photometric readings from a colorCAL colorimeter (Cambridge Research System).

In the first psychophysical experiment, to quantify crowding strength, orientation discrimination thresholds for the high- and low-contrast target gratings in both the unflanked and the flanked conditions were measured with a temporal two-alternative forced-choice (2AFC) QUEST staircase procedure (82% correct). In a discrimination

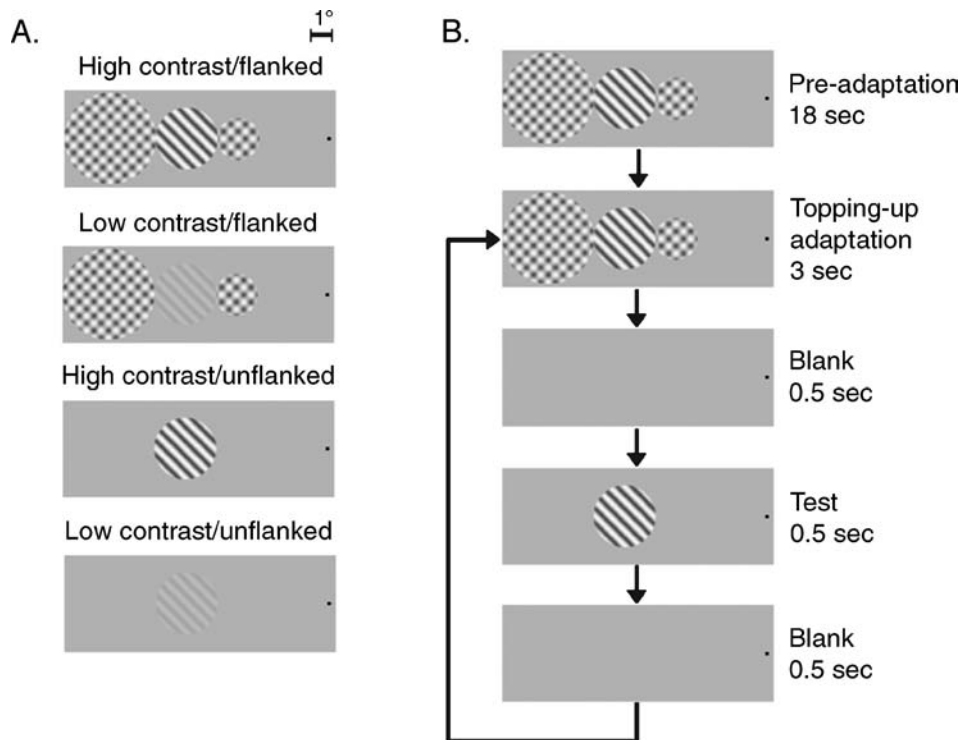


Figure 1. Stimuli and experimental design. (A) Adapting stimuli used in psychophysical and fMRI experiments. A high-contrast (0.9) or low-contrast (0.14) tilted grating as target was presented either in isolation (unflanked condition) or with two plaid flankers (flanked condition). (B) Schematic description of fMRI experiment. With 18-s pre-adaptation at the beginning of each run, a test stimulus was presented for 500 ms after 3 s topping-up adaptation. The test stimulus was either parallel or orthogonal to the adapting (target) stimulus.

trial, 45° and $45^\circ \pm \theta^\circ$ tilted target gratings were each presented for 200 ms, accompanied by an auditory beep and separated by a 600-ms blank interval. Subjects needed to indicate the rotation direction (clockwise or counterclockwise) from the first grating to the second one. The threshold was measured eight times for each condition. Crowding index was calculated as the ratio of the threshold in the flanked condition to that in the unflanked condition.

In the second psychophysical experiment, to quantify the effect of crowding on orientation-selective TEAE, contrast detection thresholds were measured with a temporal 2AFC QUEST staircase procedure (82% correct). Each staircase started with 18-s pre-adaptation. After a 3-s topping-up adaptation and a 500-ms blank interval, subjects needed to identify in which of the two successive intervals a test grating was presented. The intervals were each presented for 200 ms, accompanied by an auditory beep and separated by a 400-ms blank interval. The test grating could be either parallel or orthogonal to the adapting (target) grating. Its size, spatial frequency, and location were identical to the target grating. The test contrast varied trial by trial, controlled by QUEST. To control attention during adaptation, subjects were asked to perform a fixation task in which they needed to press one of two buttons to indicate a

250-ms luminance change (increase or decrease) of the fixation point as soon as possible. The luminance changes occurred randomly during pre-adaptation and topping-up adaptation periods, with an inter-trial interval drawn from a uniform distribution with a mean of 1.5 s and a range of 1.4–1.6 s. At the end of each topping-up adaptation period, a small arrow at the fixation point directed subjects' attention to the left visual field for detecting the test grating appeared in one of the two intervals. Contrast detection thresholds were also measured before adaptation. The experimental procedure was identical except that the pre-adaptation and the topping-up adaptation were removed. The threshold was measured eight times for each condition. TEAE was calculated as the ratio of the threshold after adaptation to that before adaptation.

The fMRI experiment consisted of 24 adaptation runs, 6 for each adapting stimulus. Each 273-s run consisted of 54 continuous trials and began with an 18-s pre-adaptation (Figure 1B). There were three types of trials—parallel, orthogonal, and blank trials. In the parallel and orthogonal trials, after a 3-s topping-up adaptation and a 0.5-s blank interval, a test stimulus was presented for 0.5 s, followed by a 0.5-s blank interval. The test stimulus was either the same as the target grating (parallel trial) or generated by rotating the target grating by 90° (orthogonal trial). The

test contrast was equal to the target contrast. In the blank trials, the 3-s topping-up adaptation was followed by a 1.5-s blank interval.

For each adapting stimulus, there were a total of 18×6 trials, 108 for each type of trial. The order of the three types of trials (parallel, orthogonal, and blank) was counterbalanced across 6 adaptation runs using M-sequences (Buracas & Boynton, 2002). These are pseudo-random sequences which have the advantage of being perfectly counterbalanced n trials back, so that each type of trial was preceded and followed equally often by all types of trials, including itself.

To control attention during adaptation, subjects performed a fixation task in which they needed to press one of two buttons to indicate a 250-ms luminance change (increase or decrease) of the fixation point as quickly as possible. The luminance changes occurred randomly through the whole run, with an inter-trial interval drawn from a uniform distribution with a mean of 1.5 s and a range of 1.4–1.6 s.

Retinotopic visual areas were defined by a standard phase-encoded method developed by Sereno et al. (1995) and Engel, Glover, and Wandell (1997), in which subjects viewed rotating wedge and expanding ring stimuli that created traveling waves of neural activity in visual cortex. Three checkered patches flickering in counterphase at 2 Hz were used in a block-design run to localize three regions of interest (ROI) corresponding to the locations of the target and distractors in the adaptation runs. The patches were at full contrast and had the same size as the target and flankers. The ROI run consisted of five cycles, and each cycle consisted of four 12-s blocks for presenting the three patches and a blank interval. This run started with a 12-s blank interval.

MRI data acquisition

In the scanner, the stimuli were back-projected via a video projector (refresh rate: 60 Hz; spatial resolution: 1024×768 ; mean luminance: 90 cd/m^2) onto a translucent screen placed inside the scanner bore. Subjects viewed the stimuli through a mirror located above their eyes. The viewing distance was 80 cm. Functional MRI data were collected using a 3-T Siemens Trio scanner with a 12-channel phase-array coil. BOLD signals were measured with an EPI sequence (TE: 30 ms; TR: 1500 ms; FOV: $196 \times 196 \text{ mm}^2$; matrix: 128×128 ; flip angle: 90; slice thickness: 2 mm; gap: 0 mm; number of slices: 21; slice orientation: axial). fMRI slices covered the occipital lobe. A high-resolution 3D structural data set (3D MPRAGE; $1 \times 1 \times 1 \text{ mm}^3$ resolution) was collected in the same session before the functional runs. All five subjects underwent five fMRI sessions, one for retinotopic mapping, two for the high contrast adaptation experiment, and two for the low contrast adaptation experiment. Total scanning time for each subject was about four hours.

MRI data processing and analysis

The anatomical volume for each subject in the retinotopic mapping session was transformed into the AC–PC space and then inflated using BrainVoyager QX. The functional volumes in all the sessions for each subject were preprocessed, which included 3D motion correction, linear trend removal, and high-pass (0.015 Hz) (Smith et al., 1999) filtering using BrainVoyager QX. Head motion within any fMRI session was less than 1.5 mm for all subjects. The images were then aligned to the anatomical volume in the retinotopic mapping session and transformed into the AC–PC space. The first 10 s of BOLD signals was discarded to minimize transient magnetic-saturation effects.

A general linear model (GLM) procedure was used for localizing ROIs in early visual areas—V1, V2, and V3. The ROIs were defined as areas that responded more strongly to the corresponding flickering circular patch than to the blank interval ($p < 10^{-4}$, corrected). We were unable to separate responses in V2 from those in V3 because our stimuli were presented close to the horizontal meridian represented at the border between these areas, and so they are treated as a single ROI. Thus, we localized six ROIs in the right hemisphere, three in V1 and the other three in V2 and V3. Even with a high statistical threshold, there were still a few voxels defined to belong to more than one ROI in some subjects. We excluded these voxels from further analyses, which meant that all the ROIs were spatially non-overlapping. From low to high eccentricity, the ROIs averaged 202, 186, and 180 mm^3 in V1 and 150, 164, and 159 mm^3 in V2/V3.

Event-related BOLD signals were calculated separately for each ROI in each subject and condition, following the method used by Larsson, Landy, and Heeger (2006), Liu, Larsson, and Carrasco (2007), and Fang, Boyaci, and Kersten (2009). For each fMRI run, the time course of MR signal intensity was first extracted by averaging the data across all the voxels within the pre-defined ROI and then normalized by the mean intensity across the run. Event-related averages were then performed for each of the three trial types (parallel, orthogonal, and blank) by averaging eight time points (12 s) starting at the test stimulus. The average response to the blank trials was subtracted from the averages of the parallel and orthogonal trials to isolate the response to the test stimulus.

The average BOLD signal at 4.5 and 6 s to the test stimulus was taken as the measure of response amplitude for each condition in subsequent analyses. Adaptation effect was quantified in two ways—amplitude difference between the parallel and orthogonal trials and adaptation index. A large positive amplitude difference (orthogonal minus parallel) means a strong adaptation effect. Adaptation index I_A was computed relative to the overall response to the stimuli in each visual area to quantify how much the measured response changed after adaptation. The index was calculated as: $I_A = (A_{\text{Orthogonal}} -$

$A_{\text{Parallel}} / (A_{\text{Orthogonal}} + A_{\text{Parallel}})$, where $A_{\text{Orthogonal}}$ is the mean amplitude of the responses to the orthogonal test stimulus and A_{Parallel} is the mean amplitude of the responses to the parallel test stimulus. This index could range from -1 to 1 and was positive whenever the mean response to the orthogonal test stimulus was greater than the mean response to the parallel test stimulus. A large adaptation index of an ROI means a strong adaptation effect.

Results

Behavioral results

Subjects' responses to the luminance change at the fixation point were both accurate and fast, indicating that they followed our instruction. Their reaction time (RT) and correct rate (CR) were categorized into several groups according to the type of adapting stimulus in the psychophysical experiment and the trial type in the fMRI experiment. There was no significant difference in RT and CR between any pair of groups in both the psychophysical and fMRI experiments. This result suggests that subjects' general attentional state did not differ across different conditions.

We estimated crowding strength by measuring orientation discrimination thresholds for a test grating in the unflanked and flanked conditions. The strength was quantified by crowding index, the ratio of the discrimination threshold in the flanked condition to that in the unflanked condition. Both indices for the high- and low-contrast target conditions were significantly larger than one (contrast 0.9: $t(4) = 42.483$, $p < 0.001$; contrast 0.14: $t(4) = 33.188$, $p < 0.001$), but there was no significant difference between them ($t(4) = 2.587$, $p = 0.061$) (Figure 2A). This result suggests that adding flankers induced a crowding effect and impaired orientation discrimination to the target, regardless of the contrast level of the target.

The effect of crowding on orientation-selective adaptation was estimated by measuring TEAEs in the unflanked and flanked conditions. Figure 2B shows that adapting to a parallel target grating significantly increased the contrast detection threshold. The TEAE from high-contrast adaptation was significantly higher than that from low-contrast adaptation (unflanked condition: $t(4) = 6.548$, $p < 0.01$; flanked condition: $t(4) = 8.091$, $p < 0.001$), which is consistent with previous reports (Blake et al., 2006; Sclar, Lennie, & DePriest, 1989). However, when subjects' attention was controlled, we failed to find significant difference between the unflanked and flanked conditions not only in the high-contrast adaptation condition ($t(4) = 1.08$, $p = 0.341$) but also in the low-contrast adaptation condition ($t(4) = 1.483$, $p = 0.212$). We also measured contrast detection thresholds after adapting to an orthogonal

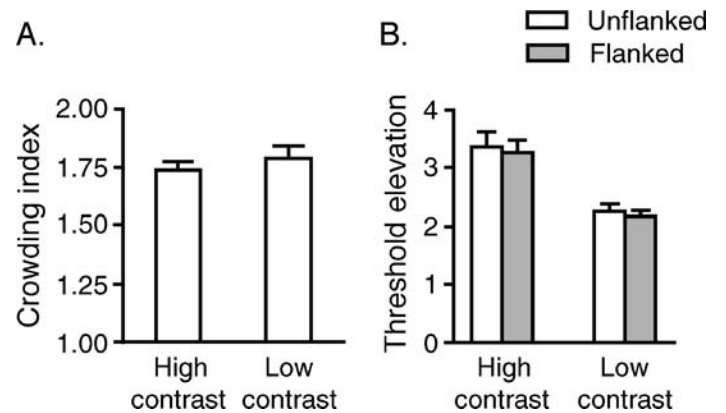


Figure 2. Psychophysical results. (A) Crowding indices for high- and low-contrast targets. Crowding index was defined as the ratio of the orientation discrimination threshold when a target was presented with flankers (flanked condition) to that when the target was presented in isolation (unflanked condition). (B) TEAEs after adapting to a high- or low-contrast target when the target was presented either in isolation (unflanked condition) or with two plaid flankers (flanked condition). TEAE is defined as the ratio of contrast detection threshold after adaptation to that before adaptation. Error bars denote 1 SEM calculated across subjects.

target grating. The TEAEs were very small, and no effect of crowding in either the high- or low-contrast adaptation conditions was found.

fMRI results

Figures 3 and 4 show the averaged fMRI results for the high- and low-contrast adaptation conditions, respectively. All subjects showed a very consistent pattern. The signals were from the ROIs in V1 and V2/V3 corresponding to the target location and peaked at 4.5 or 6 s after test stimulus onset (Figures 3A and 4A). For all adaptation conditions (high/low contrast and unflanked/flanked), fMRI adaptation effects were evident in both V1 and V2/V3—the signal in the orthogonal trials was significantly higher than that in the parallel trials (all $t(4) > 3.943$, $p < 0.017$).

It was of interest to us how crowding affected orientation-selective fMRI adaptation. fMRI adaptation effect was quantified in two ways—amplitude difference between the parallel and orthogonal trials and adaptation index (Figures 3B and 4B). In V1, no significant difference in adaptation effect was found between the unflanked and flanked conditions (high contrast/amplitude difference: $t(4) = 0.528$, $p = 0.625$; high contrast/adaptation index: $t(4) = 0.41$, $p = 0.703$; low contrast/amplitude difference: $t(4) = 0.309$, $p = 0.773$; low contrast/adaptation index: $t(4) = 0.363$, $p = 0.735$). However, in V2/V3, the adaptation effects in the unflanked condition were significantly stronger than those in the flanked condition, regardless of the adaptation contrast (high contrast/amplitude

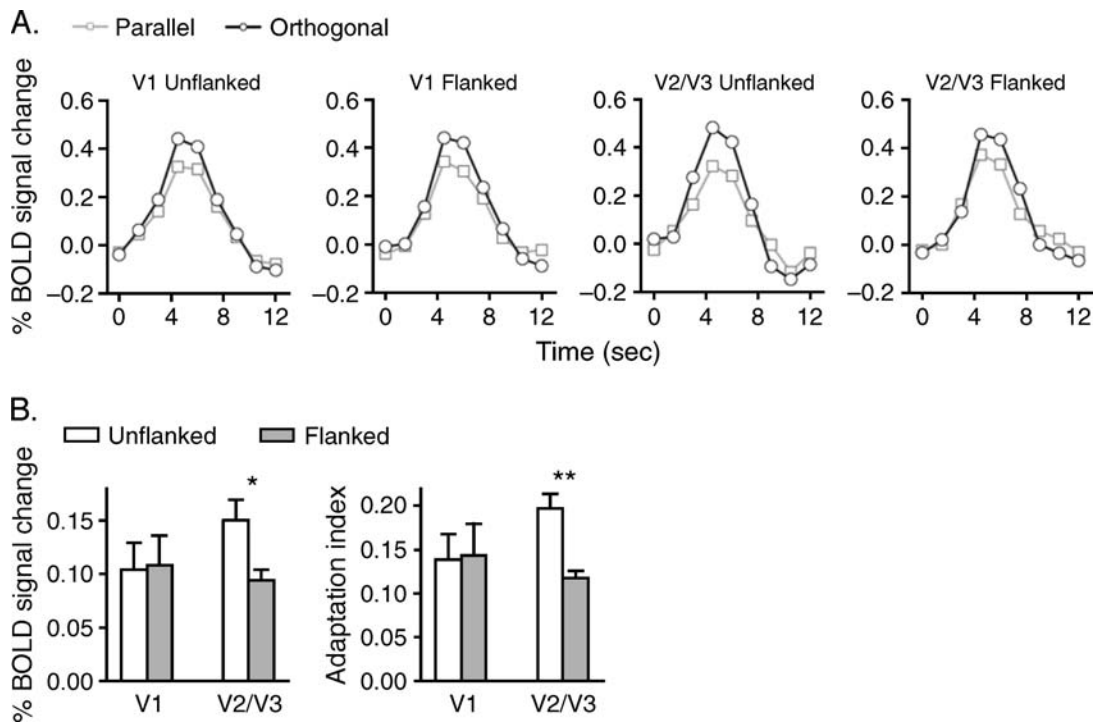


Figure 3. fMRI results for the high-contrast adaptation condition. (A) Time courses of BOLD signals in V1 and V2/V3 evoked by orthogonal and parallel test stimuli for the unflanked and flanked conditions. (B) fMRI adaptation effects in V1 and V2/V3 for the unflanked and flanked conditions as measured by the peak signal difference between the orthogonal and parallel test stimuli (left panel) and adaptation index (right panel). Asterisks indicate a statistically significant difference between the fMRI adaptation effects in the unflanked and flanked conditions (* $p < 0.05$; ** $p < 0.01$). Error bars denote 1 SEM calculated across subjects.

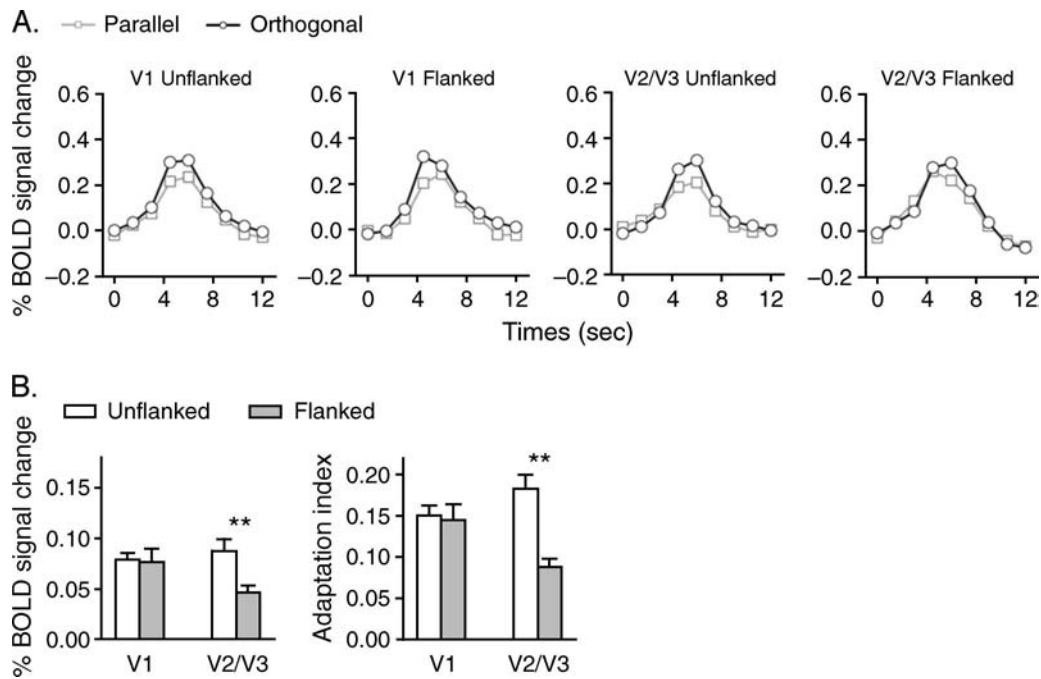


Figure 4. fMRI results for the low-contrast adaptation condition. (A) Time courses of BOLD signals in V1 and V2/V3 evoked by orthogonal and parallel test stimuli for the unflanked and flanked conditions. (B) fMRI adaptation effects in V1 and V2/V3 for the unflanked and flanked conditions as measured by the peak signal difference between the orthogonal and parallel test stimuli (left panel) and adaptation index (right panel). Asterisks indicate a statistically significant difference between the fMRI adaptation effects in the unflanked and flanked conditions (* $p < 0.05$; ** $p < 0.01$). Error bars denote 1 SEM calculated across subjects.

difference: $t(4) = 4.214$, $p < 0.05$; high contrast/adaptation index: $t(4) = 5.661$, $p < 0.01$; low contrast/amplitude difference: $t(4) = 6.633$, $p < 0.01$; low contrast/adaptation index: $t(4) = 7.297$, $p < 0.01$).

The above statistical analyses suggested that the adaptation effect in V2/V3, but not in V1, could be affected by crowding. To further examine the interaction of visual area and crowding condition, we performed two-way repeated-measures ANOVAs (visual area [V1 vs. V2/V3] \times crowding condition [unflanked vs. flanked]) for both the high- and low-contrast adaptation conditions. The interaction effects were significant for both conditions (high contrast/amplitude difference: $F(1,5) = 10.418$, $p < 0.05$; high contrast/adaptation index: $F(1,5) = 19.336$, $p < 0.05$; low contrast/amplitude difference: $F(1,5) = 10.237$, $p < 0.05$; low contrast/adaptation index: $F(1,5) = 16.192$, $p < 0.05$).

We also examined fMRI signals from the ROIs corresponding to the locations of the flankers. The signals were very weak and no adaptation effect was observed. This was not surprising because the test stimuli presented at the target location were not expected to activate these ROIs.

Discussion

We show that when subjects' attention was controlled, orientation-specific TEAE was not affected by crowding regardless of the contrast level of the adapting stimulus. The adapting contrast levels (0.9 and 0.14) here were the same as those in Blake et al. (2006; personal communication). More important, we could replicate the findings in Blake et al. if subjects were not asked to do the fixation task and were allowed to deploy their attention freely. Data from two representative subjects are presented in Figure 5. These results suggest that the reduced TEAE from adapting to a crowded low-contrast grating observed by Blake et al. might be explained by less attention to the grating.

At the cortical level, the orientation-selective adaptation in V1 was not affected by crowding for both the high- and the low-adapting contrast levels, which was parallel to the psychophysical results. This reflects a close relation between orientation-specific TEAE and orientation-selective fMRI adaptation in V1 (Fang et al., 2005), which is in line with the general belief that TEAE arises from neural activities within V1 (Movshon & Lennie, 1979). Downstream from V1, we found that crowding weakened the adaptation effect in V2 and V3. Arman et al. (2006) varied the distance between the target and flankers to manipulate the strength of the crowding effect. They found that this manipulation affected the overall V2 (but not V1) response to the target and the flankers, which is consistent with our results.

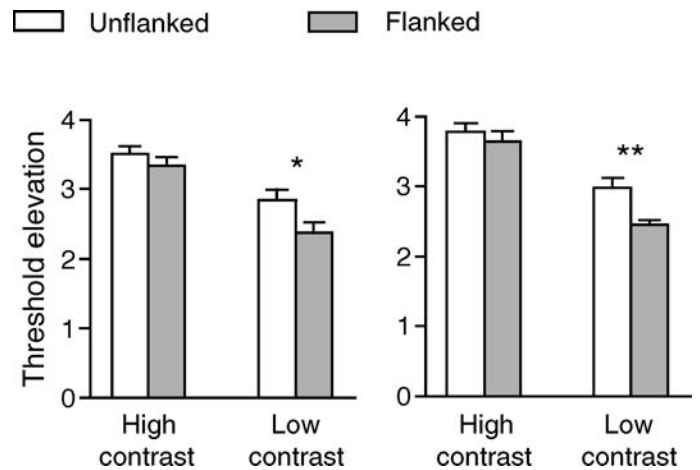


Figure 5. TEAEs from two subjects after adapting to a high- or low-contrast target in the unflanked and flanked conditions. No fixation task was needed to be performed by the subjects. Asterisks indicate a statistically significant difference between the threshold elevations in the unflanked and flanked conditions ($*p < 0.05$; $**p < 0.01$). Error bars denote 1 SE.

Since the psychophysical adaptation effect is parallel to the fMRI adaptation effect in V1, rather than V2/V3, a related question is how V1 activity can influence behavior in some way that “bypasses” V2/V3. Currently, we do not have a definite answer to this question. However, our study is not the only study showing that V1 activity, not V2/V3, is closely correlated with behavior. Similar phenomena can be found in many other fMRI studies regarding oblique effect (Furmanski & Engel, 2000), perceptual learning (Furmanski, Schluppeck, & Engel, 2004), surround suppression (Zenger-Landolt & Heeger, 2003), and bistable vision (Fang, Kersten, & Murray, 2008). A tentative answer is that the visual system might rely on the most informative cortical area to make a behavioral decision (e.g., contrast information in V1).

It should be noted that there were some differences (i.e. contrast, frame rate, stimulus duration, and luminance) between the stimuli used in the psychophysical and the fMRI experiments. Ideally, identical stimuli should be used in both experiments. Below we explain why different stimuli were used and why the differences would not affect our conclusion. Regarding the contrast difference, test stimuli with near-threshold contrast were used in the psychophysical experiment, but test stimuli with supra-threshold contrast were used in the fMRI experiment. Contrast threshold elevation with near-threshold test stimuli after adaptation is a well-established and understood phenomenon that supports the existence of orientation-tuned neurons in the visual system. Since both contrast threshold elevation and fMRI adaptation indirectly measured neural activities of different orientation-tuned neurons, we feel it is reasonable to compare the psychophysical and fMRI results. Why then did not we

use supra-threshold test stimuli in the psychophysical experiment? Blakemore, Muncey, and Ridley (1973) and Blakemore and Nachmias (1971) have shown that, after adaptation, both contrast threshold elevation with near-threshold test stimuli and loss of perceived contrast with supra-threshold test stimuli are tuned to the adapting orientation. However, the issue of whether contrast adaptation can reduce perceived contrast of test stimulus has been the subject of considerable debate (Barrett, McGraw, & Morrill, 2002; Ross & Speed, 1996; Snowden & Hammett, 1992). Contrast matching performance most likely involves mechanisms beyond V1. For example, the perceived contrast of a pair of 5 and 20 cpd gratings both at 100% contrast are quite well matched even though contrast sensitivities are very different for these two gratings (the so-called contrast constancy phenomenon), yet V1 response is much higher to the 5-cpd grating than to the 20-cpd grating (De Valois & De Valois, 1990). It seems reasonable, for this particular experiment, to use a manipulation that is well understood and more closely related to the sensitivities of the early visual cortex. Why did not we use near-threshold test stimuli in the fMRI experiment? Such a stimulus would be too weak to evoke detectable fMRI signal when subjects do not pay attention to it (Buracas & Boynton, 2007; Murray, 2008). Regarding the differences in frame rate and stimulus duration, we collected TEAE data from two subjects under the condition of 65-Hz frame rate and 500-ms test duration, which showed a very similar pattern to those under the condition of 80-Hz frame rate and 200-ms test duration. Thus, frame rate and duration did not matter in our study. Regarding the luminance difference, the monitor luminance in the psychophysical experiment was lower than the projector luminance in the fMRI experiment. To investigate if a different luminance level could change our conclusion, we re-measured the TEAE under several luminance conditions, with one luminance condition close to the projector luminance. Our data showed that the variation of luminance level did not change our conclusion.

There seems to be a growing consensus for a two-stage model of crowding (Pelli & Tillman, 2008)—feature detection, perhaps in V1 (Li, 2002), and feature integration or interaction downstream from V1. In crowding, the features in the target and flankers could be detected normally, but they are combined into a jumbled percept in the second stage. Our study suggests that crowding might be inaugurated in V2 and V3 and thus provides one of the first pieces of direct evidence supporting this model. V4 has been proposed to be a potential locus of crowding (Levi, 2008) because of its critical role in feature integration (Desimone & Duncan, 1995) and of the similarity between its receptive field size/anisotropy and the spatial extent of crowding (Pinon, Gattass, & Sousa, 1998; Toet & Levi, 1992). In this study, although we could not measure responses in more anterior visual areas beyond V3 for the reason that increasing receptive field sizes greatly reduced signal from the ROI run (Furmanski et al., 2004), it would

not be surprising to find a crowding effect in V4 simply considering its inheritance from earlier visual cortical areas.

As mentioned above, crowding occurs under a wide range of conditions and tasks. It should be noted that, although our data suggest V2/V3 as a cortical locus of crowding, it is important to test if our conclusion can be generalized to other conditions. Psychophysical evidence has suggested that crowding occurs at multiple stages in the visual system. For example, Louie et al. (2007) demonstrated a holistic crowding between high-level face representations, which suggested that face-selective areas (e.g., FFA) might play a role in this kind of crowding. In the future, complementary to psychophysical studies, more brain imaging and single-unit studies with various kinds of stimuli are needed to be carried out to obtain a full understanding of crowding—a bottleneck to object recognition.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Project 30870762), the Scientific Research Foundation for the Returned Overseas Chinese Scholars (State Education Ministry), the Ministry of Science and Technology of China (2005CB522800, 2004CB318101), and the Knowledge Innovation Program of the Chinese Academy of Sciences.

Author contributions: Taiyong Bi and Peng Cai contributed equally to this work.

Commercial relationships: none.

Corresponding author: Fang Fang.

Email: ffang@pku.edu.cn.

Address: Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, 5 Yiheyuan Road, Beijing 100871, P.R. China.

References

- Arman, A. C., Chung, S. T. L., & Tjan, B. S. (2006). Neural correlates of letter crowding in the periphery [Abstract]. *Journal of Vision*, 6(6):804, 804a, <http://journalofvision.org/6/6/804/>, doi:10.1167/6.6.804.
- Barrett, B. T., McGraw, P. V., & Morrill, P. (2002). Perceived contrast following adaptation: The role of adapting stimulus visibility. *Spatial Vision*, 16, 5–19. [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]
- Blakemore, C., Muncey, J. P., & Ridley, R. M. (1973). Stimulus specificity in the human visual system. *Vision Research*, 13, 1915–1931. [PubMed]

- Blakemore, C., & Nachmias, J. (1971). The orientation specificity of two visual after-effects. *The Journal of Physiology*, *213*, 157–174. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178. [PubMed]
- Buracas, G. T., & Boynton, G. M. (2002). Efficient design of event-related fMRI experiments using M-sequences. *Neuroimage*, *15*, 801–813. [PubMed]
- Buracas, G. T., & Boynton, G. M. (2007). The effect of spatial attention on contrast response functions in human visual cortex. *Journal of Neuroscience*, *27*, 93–97. [PubMed] [Article]
- Chung, S. T. L., Li, R. W., & Levi, D. M. (2007). Crowding between first- and second-order letter stimuli in normal foveal and peripheral vision. *Journal of Vision*, *7*(2):10, 1–13, <http://journalofvision.org/7/2/10/>, doi:10.1167/7.2.10. [PubMed] [Article]
- De Valois, R. L., & De Valois, K. K. (1990). *Spatial vision*. Oxford: Oxford University Press.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. [PubMed]
- Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*, 659–671. [PubMed]
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, *7*, 181–192. [PubMed]
- Fang, F., Boyaci, H., & Kersten, D. (2009). Border ownership selectivity in human early visual cortex and its modulation by attention. *Journal of Neuroscience*, *29*, 460–465. [PubMed] [Article]
- Fang, F., & He, S. (2008). Crowding alters the spatial distribution of attention modulation in human primary visual cortex. *Journal of Vision*, *8*(9):6, 1–9, <http://journalofvision.org/8/9/6/>, doi:10.1167/8.9.6. [PubMed] [Article]
- Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision*, *8*(7):2, 1–9, <http://journalofvision.org/8/7/2/>, doi:10.1167/8.7.2. [PubMed] [Article]
- Fang, F., Murray, S. O., Kersten, D. J., & He, S. (2005). Orientation-tuned fMRI adaptation in human visual cortex. *Journal of Neurophysiology*, *94*, 4188–4195. [PubMed] [Article]
- Festman, Y., & Ahissar, M. (2004). Attentional states and the degree of visual adaptation to gratings. *Neural Networks*, *17*, 849–860. [PubMed]
- Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contour interaction and visual resolution: Contralateral effect. *Science*, *142*, 979–980. [PubMed]
- Furmanski, C. S., & Engel, S. A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, *3*, 535–536. [PubMed]
- Furmanski, C. S., Schluppeck, D., & Engel, S. A. (2004). Learning strengthens the response of primary visual cortex to simple patterns. *Current Biology*, *14*, 573–578. [PubMed]
- Gardner, J. L., Sun, P., Waggoner, R. A., Ueno, K., Tanaka, K., & Cheng, K. (2005). Contrast adaptation and representation in human early visual cortex. *Neuron*, *47*, 607–620. [PubMed] [Article]
- Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Reviews*, *78*, 467–485. [PubMed] [Article]
- He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*, 334–337. [PubMed]
- He, S., Cavanagh, P., & Intriligator, J. (1997). Attentional resolution. *Trends in Cognitive Sciences*, *1*, 115–121.
- Hess, R. F., Dakin, S. C., & Kapoor, N. (2000). The foveal “crowding” effect: Physics or physiology? *Vision Research*, *40*, 365–370. [PubMed]
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, *43*, 171–216. [PubMed]
- Korte, W. (1923). Über die Gestaltauffassung im indirekten Sehen. *Zeitschrift für Psychologie*, *93*, 17–82.
- 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. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*, 635–654. [PubMed]
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, *25*, 963–977. [PubMed]
- Li, X., Lu, Z., Tjan, B. S., Doshier, B. A., & Chu, W. (2008). Blood oxygenation level-dependent contrast response functions identify mechanisms of covert attention in early visual areas. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 6202–6207.
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*, 9–16. [PubMed]
- Liu, L., & Arditi, A. (2000). Apparent string shortening concomitant with letter crowding. *Vision Research*, *40*, 1059–1067. [PubMed]

- Liu, T., Jiang, Y., Sun, X., & He, S. (2009). Reduction of the crowding effect in spatially adjacent but cortically remote visual stimuli. *Current Biology*, *19*, 127–132.
- Liu, T., Larsson, J., & Carrasco, M. (2007). Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron*, *55*, 313–323. [PubMed]
- Louie, G. E., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision*, *7*(2):24, 1–11, <http://journalofvision.org/7/2/24/>, doi:10.1167/7.2.24. [PubMed] [Article]
- Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, *5*(1):6, 58–70, <http://journalofvision.org/5/1/6/>, doi:10.1167/5.1.6. [PubMed] [Article]
- Movshon, J. A., & Lennie, P. (1979). Pattern selective adaptation in striate cortical neurones. *Nature*, *278*, 850–852.
- Murray, S. O. (2008). The effects of spatial attention in early human visual cortex are stimulus independent. *Journal of Vision*, *8*(10):2, 1–11, <http://journalofvision.org/8/10/2/>, doi:10.1167/8.10.2. [PubMed] [Article]
- Nandy, A. S., & Tjan, B. S. (2007). The nature of letter crowding as revealed by first- and second-order classification images. *Journal of Vision*, *7*(2):5, 1–26, <http://journalofvision.org/7/2/5/>, doi:10.1167/7.2.5. [PubMed] [Article]
- 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., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*, 1129–1135. [PubMed]
- Pinon, M. C., Gattass, R., & Sousa, A. P. (1998). Area V4 in Cebus monkey: Extent and visuotopic organization. *Cerebral Cortex*, *8*, 685–701. [PubMed]
- Ross, J., & Speed, H. D. (1996). Perceived contrast following adaptation to gratings of different orientations. *Vision Research*, *36*, 1811–1818. [PubMed]
- Scialo, G., Lennie, P., & DePriest, D. D. (1989). Contrast adaptation in striate cortex of macaque. *Vision Research*, *29*, 747–755. [PubMed]
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893. [PubMed]
- Smith, A. M., Lewis, B. K., Ruttimann, U. E., Ye, F. Q., Sinnwell, T. M., Yang, Y., et al. (1999). Investigation of low frequency drift in fMRI signal. *Neuroimage*, *9*, 526–533. [PubMed]
- Snowden, R. J., & Hammett, S. T. (1992). Subtractive and divisive adaptation in the human visual system. *Nature*, *355*, 248–250. [PubMed]
- Spivey, M. J., & Spirm, M. J. (2000). Selective visual attention modulates the direct tilt aftereffect. *Perception & Psychophysics*, *62*, 1525–1533. [PubMed]
- 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]
- Strasburger, H., Harvey, L. O. J., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics*, *49*, 495–508. [PubMed]
- Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty: Its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology*, *53*, 471–477. [PubMed]
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*, 1349–1357. [PubMed]
- Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research*, *42*, 2357–2369.
- Westheimer, G., Shimamura, K., & McKee, S. P. (1976). Interference with line-orientation sensitivity. *Journal of the Optical Society of America*, *66*, 332–338. [PubMed]
- Westheimer, G., & Truong, T. T. (1988). Target crowding in foveal and peripheral stereoacuity. *American Journal of Optometry and Physiological Optics*, *65*, 395–399. [PubMed]
- Wilkinson, F., Wilson, H. R., & Ellernberg, D. (1997). Lateral interactions in peripherally viewed texture array. *Journal of the Optical Society of America*, *14*, 2057–2068.
- Williams, R. A., & Essock, R. A. (1986). Areas of spatial interaction for a hyperacuity stimulus. *Vision Research*, *26*, 349–360.
- Zenger-Landolt, B., & Heeger, D. J. (2003). Response suppression in V1 agrees with psychophysics of surround masking. *Journal of Neuroscience*, *23*, 6884–6893.