

Legibility of Chinese characters in peripheral vision and the top-down influences on crowding

Jin-Yun Zhang^{a,1}, Ting Zhang^{a,1}, Feng Xie^b, Lei Li^{c,*}, Cong Yua^{*,*}

^aSa e Ke Lab a C e e Ne ce ce a d Lea e , Be N a U e e , Be 100875, C a

^bEENT H a, F da U e e , S a a, C a

^cSc O e e, U e e Aaba a a B e a , B e a , AL, USA

ARTICLE INFO

Accepted:

Received 5 March 2008

Received in revised form 24 September 2008

Keywords:

Chinese characters

Legibility

Crowding

Top-down influences

Peripheral vision

ABSTRACT

Written Chinese is disincarnated from alphabetic languages because of its enormous number of characters with a great range of spatial complexities (stroke numbers). In this study, we investigated the impact of spatial complexity on legibility of Chinese characters as well as associated crowding in peripheral vision. Our results showed that for isolated characters, threshold sizes of complex characters increased faster with relative eccentricity than did those of simple characters, suggesting possible "within-character crowding" among parts of complex Chinese characters. However, such "within-character crowding" was rendered negligible by strong between-character crowding in rod-and-cone backgrounds. When the large and small characters belonged to different complexity groups, the intensity and extent of crowding were greatly reduced, which could be explained by top-down influences as well as lower-level mechanisms. We suggest that crowding can be attributed to multiple mechanisms at different levels of visual processing.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Most studies of letter legibility use Roman letters. Roman letters are highly similar to each other because they are made of a small number of strokes, have no discernible parts, and are relatively uniform in spatial complexity as a stimulus set. It is less clear how much of our knowledge obtained from such stimuli can be applied to legibility of Chinese characters (CCs) that contain 1 to as many as 52 strokes, and thus have a wide range of spatial complexities. Recently, we reported a study on legibility of CCs in foveal vision (Zhang, Zhang, Xie, Li, & Yu, 2007), in which we measured threshold (acuity) sizes for six groups of frequently used CCs from low to high spatial complexity, and determined the relationship between legibility and optical defocus for Landolt C, Snellen E and three groups of CCs representing low, medium, and high spatial complexities. Our results showed that CC acuity sizes increase steadily with stimulus complexity, although a slower rate than would be expected if visual acuity is based on discerning the fine details of the stimulus. Moreover, the acuity sizes of optical defocus functions of the three CC groups and Snellen E have similar slopes, differing only by a vertical shift (approximately one, two, and three lines above E acuity on an acuity chart, respectively), suggesting the feasibility of using Snellen E acuity, which

is the current standard approach for acuity testing in China, to derive the legibility of CCs in foveal vision. To understand the slower rate of acuity size increase against spatial complexity, we also developed a geometric moment model, in which we propose that human letter recognition performance near the acuity limit can be accounted for by a set of global features described by zero-order, first-order, and perceptually meaningful low-order geometric moments (i.e., the ink area, variance, skewness, and kurtosis; Martin & Scurry, 1999).

The current study extends our work on the legibility of CCs, as well as crowding, in peripheral vision. We are particularly interested in understanding how the characteristics of CCs that would affect peripheral character legibility and crowding in a way not normally identified when alphabetic stimuli are used. First, the majority of CCs are spatially complex. Only 4% of CCs are single-bodied characters (e.g., same square area as the single-bodied CCs). We speculate that interactions among these parts could interfere with the recognition of a complex CC as a whole, and such interactions, or "within-character crowding", could be magnified in the periphery. If this is indeed the case, acuity tests of different complexity CC groups may have different spatial scaling functions in the periphery, and thus may not be derived from a standard measure like E acuity, as we have shown previously for foveal vision (Zhang et al., 2007), which

proper compensations of scaling differences among CC groups. Such a possibility would have important clinical implications in evaluating peripheral vision of patients who read e-ink characters of different spatial complexities.

To address this issue, in the first part of the study, we measured threshold sizes of single CCs of various complexities at different eccentricities. By comparing the slopes of spatial scaling functions for different complexity CCs groups, we revealed an inferiority of complex CCs to simple CCs in the visual periphery, possibly indicating "within-character crowding among parts of complex CCs. We also measured threshold sizes of masked CCs in a bigram configuration to assess the impact of within-character crowding on regular "between-character crowding.

The second distinctive characteristic of CCs we are particularly interested in is that, in real-world Chinese e-ink, more than one is a character masked by characters of different spatial complexities. Such configurations are rarely seen in alphabetic languages because alphabetic letters tend to have similar spatial complexities. In cases where the large and small characters have different spatial complexities, some basic similarity properties, such as the brightness and the spatial frequency contents, are different between the large and small characters. These and other physical similarity differences including shape, size, polarity, etc., are known to affect crowding by segregating the large and small characters (Chung, Leung, & Legge, 2001; Hess, Dakin, & Kapoor, 2000; Kooi, Toet, Tripathi, & Leung, 1994; Na, 1992). Moreover, a Chinese reader knows that large and small characters often have different spatial complexities in a bigram configuration, such as 个需十, are drawn from different similarity groups, so that he or she will not report a small character as the large one. There is evidence that such misreporting can be attributed to crowding (Srasberger, 2005). Therefore, both similarity differences and high-level top-down influences may affect crowding when the large and small characters differ in complexity.

In the second part of the study we assessed the impact of large character complexity on crowding. We also designed experiments to isolate the top-down influence on crowding, using not only CCs but also English Sloan letters. Moreover, after isolation of top-down influences, we were able to manipulate similarity features to identify lower-level mechanisms underlying crowding. On the basis of our results, as well as previously reported findings, we propose an eclectic idea that uses multiple mechanisms and multiple processing levels to explain crowding.

2. Methods

2.1. Observer adaptation

Six observers with normal or corrected-to-normal vision participated in the study. All observers were young (mean age = 23.3 years) native Chinese speakers with college education and at least 6 years of training in reading and writing English. Observers ZJ and ZT were cochlear and were experienced in psychophysical experiments. The others were new psychophysical observers and were naive to the purposes of the study. Written informed consent was obtained from all observers prior to the tests.

The stimuli were generated by a Matlab-based WinVis program (Neumeister, Insley, Oakland, CA) and were presented on a 21-in. Sony G520 color monitor (2048 pixel × 1536 pixel, 0.189 mm × 0.189 mm per pixel, 75 Hz frame rate). The minimal and maximal luminance of the monitor was 1.18 and 91 cd/m², respectively. Viewing was monocular in a dimly lit room. A head-and-chin rest was used to stabilize the head position.

2.2. Stimuli

The stimuli were generated by a Matlab-based WinVis program. The letters and four groups of Chinese characters (CC1, CC2, CC3, CC4, CC5, CC6) were presented in Sloan letters or Chinese characters (CC1, CC2, CC3, CC4, CC5, CC6) in two sizes (large and small) and two orientations (horizontal and vertical). The large characters were selected from the Sloan letters or Chinese characters (CC1, CC2, CC3, CC4, CC5, CC6) in two sizes (large and small) and two orientations (horizontal and vertical). The large characters were selected from the Sloan letters or Chinese characters (CC1, CC2, CC3, CC4, CC5, CC6) in two sizes (large and small) and two orientations (horizontal and vertical). The large characters were selected from the Sloan letters or Chinese characters (CC1, CC2, CC3, CC4, CC5, CC6) in two sizes (large and small) and two orientations (horizontal and vertical).

The spatial complexity of the stimuli was also described by stroke frequency (Zhang et al., 2007). Each letter or character was sliced at 6 directions (horizontal, vertical, and four oblique at 45° and 135°). From each slicing, the average stroke frequency was calculated. The average stroke frequency of the Sloan letters was 2.0 strokes/letter. The average stroke frequency for the six groups of CCs increased monotonically from 2.0 to 5.5 strokes/character (Zhang et al., 2007).

2.3. Procedure

The large characters were presented either in Sloan letters or Chinese characters on a full-screen format. The characters were presented either in Sloan letters or Chinese characters on a full-screen format. The characters were presented either in Sloan letters or Chinese characters on a full-screen format.

Sloan

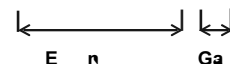
CC1 个 么 少 十

CC3 条 名 电

CC4 益 售

CC5

CC6



differ from each other and from the average. The flankers all always had the same size as the average, and the edge-to-edge average flanker gap was one character wide if inspected (Fig. 1b). The average was presented at 0°, 5°, or 10° retinal eccentricities on the horizontal meridian in the temporal visual field. The viewing distance was 6, 1.6, and 0.8 m for 0°, 5°, and 10° retinal eccentricity, respectively.

In each trial of focal testing, a 0.1° square area was displayed for 200 ms at the center of the screen accompanied by a beep, which was followed by a 300 ms time gap prior to the onset of the stimuli. The stimulus duration was 200 ms. When flankers were selected their display was asynchronous with the average. The stimuli had the same abrupt onset and offset. For peripheral testing, the central fixation was always present, and the observer was asked to read a digit. At the beginning of each trial, a small square (0.1°) was displayed for 200 ms at the average location as a location cue, which was followed by a 300 ms gap prior to the onset of the stimuli. The stimuli were presented for 200 ms. The observer was asked to identify the average from a list of the members of the average group (the list was printed on paper for observer's reference), and to report the result by pressing a number key. An auditory feedback was provided upon an incorrect response.

The threshold level is either horizontal or vertical flankers as measured in the method of constant stimuli. In Experiment I and II, which were run together, each experiment session was composed of threshold size measurement with a combination of stimuli groups, retinal eccentricities, and flanking conditions. Each threshold measurement was based on the levels of stimuli size with 10 presentations at each level. A typical round of experiments consisted of 30 sessions (5 stimulus groups \times 3 eccentricities \times 2 flanking conditions), which were run according to a random permutation for each observer and were completed in about 10 days. Each observer completed 7 rounds of the experiments. All conditions in each session were run in Experiment III and IV could be completed within a 2-h session and were repeated in several days. The percent correct data were fitted with a Weibull function: $P = 1 - (1 - \gamma)e^{-(\lambda/x)^\beta}$, where P was the percent correct, γ was the guessing rate (0.2 in a 5AFC trial), λ was the stimulus angular size, β was the slope of the psychometric function, and λ was the threshold size for recognition at a 70.6% correct level.

3. Results

3.1. Experiment I: Level, CC, eccentricity, and edge-to-edge

This experiment measured threshold sizes for four groups of isolated CCs as well as Sloan letters at 0°, 5°, and 10° retinal eccentricities. Individual and mean threshold sizes plotted against eccentricity, along with regression lines (with error bars), were shown in Fig. 2a and b. A repeated measures ANOVA indicated that for all stimulus groups, the threshold sizes increased with the retinal eccentricity linearly ($p < .001$; Fig. 2a and b). The threshold sizes of the more complex CCs (CC4 and CC6) were similar ($p = .978$), and were significantly larger than those of simpler CC1 ($p = .002$) and CC3 ($p = .026$). CC3 threshold sizes were larger than those of CC1 ($p = .032$), and CC1 threshold sizes were larger than those of Sloan letters ($p = .022$). The latter could be explained by the flicker strokes of the Sloan letters (Zhang et al., 2007).

There was a significant interaction between stimulus groups and eccentricities ($p < .001$), suggesting that the increase of threshold sizes with the retinal eccentricity was affected by the stimulus groups. To characterize this interaction, peripheral threshold sizes were normalized by corresponding focal threshold sizes. The resultant size scaling functions were shown in Fig. 2c, and the function slopes were plotted against stroke frequency in Fig. 2d. These plots showed a systematic increase of scaling function slope

from simple to more complex CCs. The slopes of CC6 and CC4 were 24% and 26% greater than those of CC3, respectively, and 56% and 59% greater than those of CC1, respectively. Moreover, when slopes of the scaling functions for four CC groups were plotted against the stimulus complexities (stroke frequencies), the slope of the regression line was significantly different from zero ($p = .002$) (Fig. 2d). These data indicated that the threshold sizes of more complex CCs (CC4 and CC6) increased a faster rate with the retinal eccentricity than did those of simpler CCs. We interpreted this systematic change of regression slope as evidence for possible interactions among components of more complex CCs, or "intra-character coding," in the visual periphery (see Section 4).

3.2. Experiment II: Central versus peripheral CC eccentricity

A letter is more difficult to identify when it is closely flanked by additional letters (Flom, Heath, & Takahashi, 1963; Schar & Brian, 1962. See Levi (2008) for a more recent review). Would such crowding between the average and flanker characters be affected by intra-average coding? In this experiment we measured the threshold sizes for flanked Sloan, CC1, CC3, CC4, and CC6 at 0°, 5°, and 10° retinal eccentricities. The average and flankers were drawn from the same 5-member stimulus group (Fig. 1a), and the edge-to-edge gap between average and flankers was always one character width (Fig. 1b). This experiment was run together with Experiment I on the same observers (see Section 2). Individual data, their averages, and the regression lines are shown in Fig. 3a and b.

As expected, strong crowding was evident in recognition of flanked Sloan letters and CCs in peripheral vision. The slopes of spatial scaling functions were much steeper for flanked average sizes (Fig. 3c, dashed lines) than for isolated average sizes (Fig. 3c, solid lines, replotted from Fig. 2c). In the focal area, threshold sizes under the flanker and no-flanker conditions were not significantly different ($p = .591$), consistent with Flom (1991) who found crowding did not extend beyond one character width.

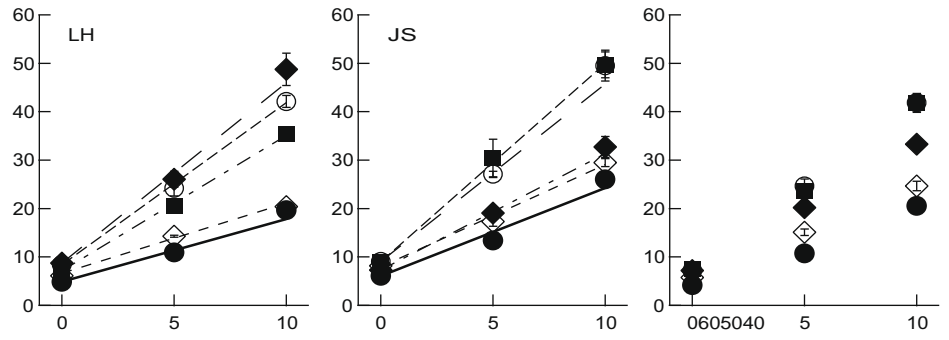
The best fitting lines of the threshold sizes versus retinal eccentricity functions became steeper with increasing CC complexity (Fig. 3a and b). However, his increase only reflected focal threshold size differences among the CC groups. When peripheral threshold sizes were normalized by corresponding focal threshold sizes, the differences among the scaling function slopes of various CC groups were insignificant ($p = .344$; Fig. 3c). When the slopes of the scaling functions for the four CC groups were plotted against stroke frequencies, the slope of the regression line was not significantly different from zero ($p = .679$) (Fig. 3d). These results suggested that when flankers were present, characters of different spatial complexities scaled in a similar manner with the retinal eccentricity.

It is important to distinguish the normalized spatial scaling factors for focal thresholds in order to distinguish from Boima (1970) non-normalized spatial scaling factors. Boima (1970) reported that the non-normalized scaling factor for critical crowding was approximately 0.5 (i.e., half the retinal eccentricity). This factor varied from 0.23 (Sloan) to 0.37 (CC6) in order that when the sizes of the critical ones were calculated in average flanker center-to-center distance at a 70.6% correct rate (the threshold sizes were in edge-to-edge gap size in Fig. 3), smaller than Boima's factor of 0.5. This difference could be due to the different criteria used to define the thresholds (Levi, 2008).

3.3. Experiment III: The effect of average eccentricity on central crowding

In the introduction it was suggested that in normal Chinese a character is more likely to have neighboring characters with different spatial complexities. Such complexity differences could

Sloan
 CC1
 CC3
 CC4
 CC6



CC4
 Sloan
 CC6
 CC3
 CC1

Fig. 3. Crossing threshold Sloan letters and Chinese characters. (a and b) Individual and mean threshold sizes as a function of retinal eccentricity. (c) Scaling factors

in rod ce lo -le el brigh ness and spial freq enc differences be een he arge and ankers. I o ld also in rod ce a op-do n in ence o segrega e he arge and ankers, especial hen he comple i difference is large. In his e perimen , e meas red he effec s of arge anker comple i con ras on cro ding i h CCs. La er in E perimen IV e o ld isola e he op-do n in ences on cro ding sing CCs as ell as English Sloan le ers as s im li.

3.3.1. T e e c a e a e c e , c a c d

To ma imi e comple i con ras , he leas and mos comple CCs, CC1 and CC6 , ere sed as arge and anker s im li. The a erage s roke freq encies ere 2.22 and 5.52 s rokes per charac- er for CC1 and CC6 s im li, respec i el . Threshold si es ere meas red a 10° re inal eccen rici for CC1 and CC6 arge s i h hree arge anker comple i con ras condi ons: (1) ero comple i con ras : a CC1 or CC6 arge i h ankers from he same 5-member s im l s gro p (deno ed as “111” and “666” condi ons. Digi s “1” and “6” s and for CC1 and CC6 charac ers, respec i el , and he lef , cen er, and righ digi s represen he lef anker, cen er arge , and righ anker, respec i el); (2) f ll comple i con ras : a CC1 arge i h CC6 ankers (“616” condi on) or a CC6 arge i h CC1 ankers (“161” condi on); (3) mi ed comple i con ras : a CC1 arge i h a CC6 anker and a CC1 anker (“611/116” condi ons) or a CC6 arge i h a CC1 anker and a CC6 anker (“166/661” condi ons). Threshold si es for single CC1 and CC6 i ho ankers ere also meas red as baselines (deno ed as “1” and “6”).

Fig. 4 sho s he hreshold si es ob ained nder ario s arge anker comple i con ras condi ons. When he arge and ankers had f ll comple i con ras s (616 and 161), cro ding as red ced signi can l from ha a ero comple i con ras (111 and 666) (= .001, repea ed meas res ANOVA), b 55.5 % 4.4% for he CC1 arge (Fig. 4, gra bars) and 34.0 % 4.2% for he CC6 arge (Fig. 4, black bars). Cro ding as red ced more for he CC1 arge b he CC6 ankers in he 616 con g ra ion han for he CC6 arge b he CC1 ankers in he 161 con g ra ion. This as mme r co ld be d e o he fac ha for he 616 con g ra ion, hen he CC1 arge as near hreshold, he CC6 ankers ere mos likel belo

heir non- anker “6” baseline hresholds (Fig. 4). Therefore, he fea res of hese CC6 ankers ere no er legible and had less chance o be improperl in egra ed i h fea res of he CC1 arge o prod ce cro ding. Ho e er, cro ding as no comple el elimina ed a f ll comple i con ras . Threshold si es for 616 and 161 condi ons ere sill signi can l larger han “1” and “6” baselines (= .002), hich ere 29.6 % 4.0% and 38.7 % 10.0% larger, respec i el .

A mi ed comple i con ras s, here as no signi can difference he her he same-gro p anker as on he lef or righ side of he arge , so he res l s ere a eraged. Cro ding a mi ed comple i con ras s (116/611 and 166/661) as eaker han ha a ero comple i con ras s (111 and 666) (= .008 and .021, respec i el , Fig. 4), b s ronger han ha a f ll comple i con ras s (616 and 161) (= .063 and .021, respec i el , Fig. 4).

Ho e er, i is or h men ioning ha he abo e es ima ion of he comple i con ras effec s ere mos conser aie , i h he ass mp ion ha he g essing ra e of he cen er arge as n- changed across ario s anker condi ons. Ho e er, le ers a he beginning and end of a le er s ring are kno n o be more legible han le ers in he middle (Wolford & Hollings or h, 1974), so i as likel ha a some charac er si es in o r e perimen s, he obser ers co ld recogni e one or boh ankers b no he arge . When boh ankers ere recogni ed, he arge g essing ra e as 1/3 nder ero comple i con ras condi ons (111 and 666) beca se boh ankers ere member of he 5-charac er s im l s gro p, and 1/5 nder f ll comple i con ras condi ons (161 and 616) beca se boh ankers ere from a differen s im l s gro p. The higher ra es of correc g essing associa ed i h he ero comple i con ras s o ld ha e ca sed nderes ima ion of he hreshold si es for he 111 and 666 condi ons, and nderes ima ion of he hreshold differences be een he ero- and f ll-comple i con ras condi ons.

3.3.2. T e e c a e a e c e , c a c , ca ac

Besides he hreshold change, cro ding is also q an i ed b i s spial e en or cri cal spacing (he one i hin hich ankers in erfere i h he arge recogni on). Se eral s dies repor ed ha

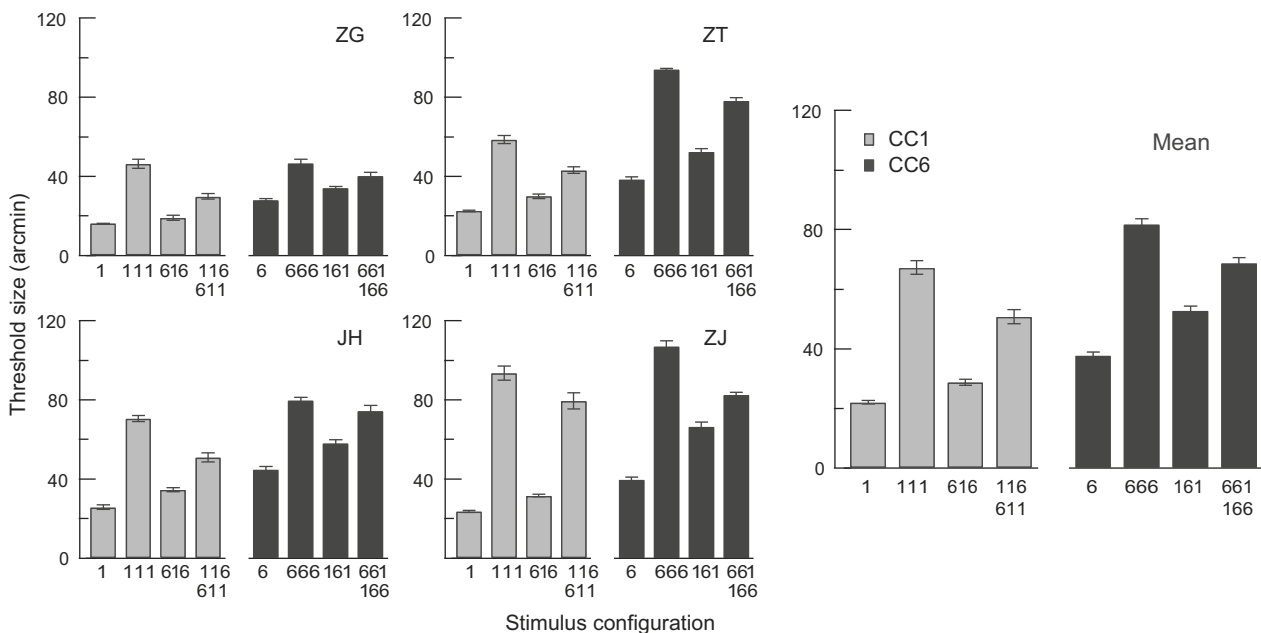


Fig. 4. The effec s of arge anker comple i con ras on cro ding. 111 and 666: ero comple i con ras s; 616 and 161: f ll comple i con ras s; 116/611 and 661/166: mi ed comple i con ras s. Digi s “1” and “6” s and for CC1 and CC6 s im li, respec i el . The lef , cen er, and righ digi s represen he lef anker, cen er arge , and righ anker, respec i el .

the critical spacing is approximately half the average retinal eccentricity regardless of the average size (Bomba, 1970; Ching et al., 2001; Pelli, Palomares, & Majaj, 2004; Tripathi & Casanovi, 2002), but the actual value depends on how the spacing is defined (center-to-center or edge-to-edge) and how the criterion is defined in terms of the crowding angle (Lee, 2008).

We measured critical spacing of crowding angle for complex patterns (111 and 666) and full complex patterns (616 and 161) at 5° and 10° retinal eccentricities for the same four observers. Critical spacing for Sloan letters and complex patterns was also measured for comparison. The sizes of the average and flankers were fixed at 1.2 times each observer's single character threshold sizes (Fig. 4), and the average correction rate was measured as a function of the average flanker center-to-center separation. Critical spacing was defined as the center-to-center separation at a 70.6% correction rate. Critical spacing for complex patterns and Sloan letters (111, 666 and SSS for Sloan letters) was statistically similar at 1.80 ± 0.47°, 2.26 ± 0.49°, and 1.85 ± 0.47° at 5° eccentricity (Fig. 5a), respectively, and at 3.17 ± 0.13°, 3.24 ± 0.44°, and 3.26 ± 0.17° at 10° eccentricity (Fig. 5b), respectively ($F = .462$, repeated measures ANOVA). However, critical spacing was significantly smaller than the average and flankers were a full complex pattern ($F = .006$), with an overall reduction of 41.0%. The 616 complex pattern condition reduced more crowding from the 111 condition (by 49.4%, averaged over 5° and 10° data, Fig. 5a and b, gray bars) than did the 161 complex pattern condition from the 666 condition (by 32.6%, averaged over 5° and 10° data, Fig. 5a and b, black bars) ($F = .006$). The reductions of critical spacing were similar at 5° and 10° retinal eccentricities ($F = .161$).

3.4. Experiment IV: The effect of crowding angle on the crowding angle

Strauss (2005) reported that under crowding an observer might report the flanking letters as the average, which was supported by error analysis using the 111 and 666 data in Fig. 4. Specifically, for all similar sizes producing less than 60% correction rate (mean = 38.6% and 37.8% for 111 and 666 conditions, respectively), the rate that the observers mistakenly reported one of the flanking characters as the average was significantly higher than the rate of reporting the other non-target characters (52.5% s.d. 8.9% for the 111 condition and 44.6% s.d. 17.6% for the 666 condition; $F < .001$, repeated measures ANOVA). These misreporting rates were calculated against the total number of included trials, not the number of wrong report trials, so the observers often reported the flankers more frequently than the correct average. However, when the average and flankers were drawn from different similar groups (i.e., 161 and 616 conditions), the observer could not report the flankers as the average, because he or she knew that the flanking characters were not one of the list of reportable characters. Besides similar size differences (i.e., brightness, spatial frequency) that might have segregated the average and flankers, how much of the crowding angle in the center of the crowding reduction in Fig. 4? In his experiment we attempted to isolate his crowding angle in the center of the crowding, as well as the underlying mechanisms that also affect crowding.

3.4.1. The effect of crowding angle on the crowding angle

To isolate the high-level crowding angle in the center, we compared crowding when the average and flankers were drawn either from the same similar group, or from different similar groups, while keeping the average flanker complex pattern constant. To make this possible, as shown in Fig. 6a, the average in the rigram was also drawn from the CC1 characters used in the previous experiment, and the flankers were either drawn from the remaining four characters ("same flanker condition in Fig. 6), or from the other char-

acters ("different flanker condition in Fig. 6). These new characters and the existing characters had similar number of strokes (2–4) and similar bigram Euclidean distances among each other (Zhang et al., 2007). Therefore, the average flanker complex patterns were either "same" and "different flanker conditions, but the flankers in the "same condition were reportable characters and the flankers in the "different conditions were not. The observers were clearly informed whether the average and flanking characters were from the same similar group or from different groups, and the similarity was listed on paper as a response guide. This design isolated the observer's knowledge of average and flanker identities as a crowding angle in the center of the crowding and controlled the impact of low-level similarity factors. We also ran a parallel experiment using Sloan letters following the same procedure. The average was drawn from the Sloan letters (CDKNS) used in the previous experiment, and the flankers were drawn either from the remaining four letters, or from the other previous Sloan letters (VROHZ, Fig. 6a).

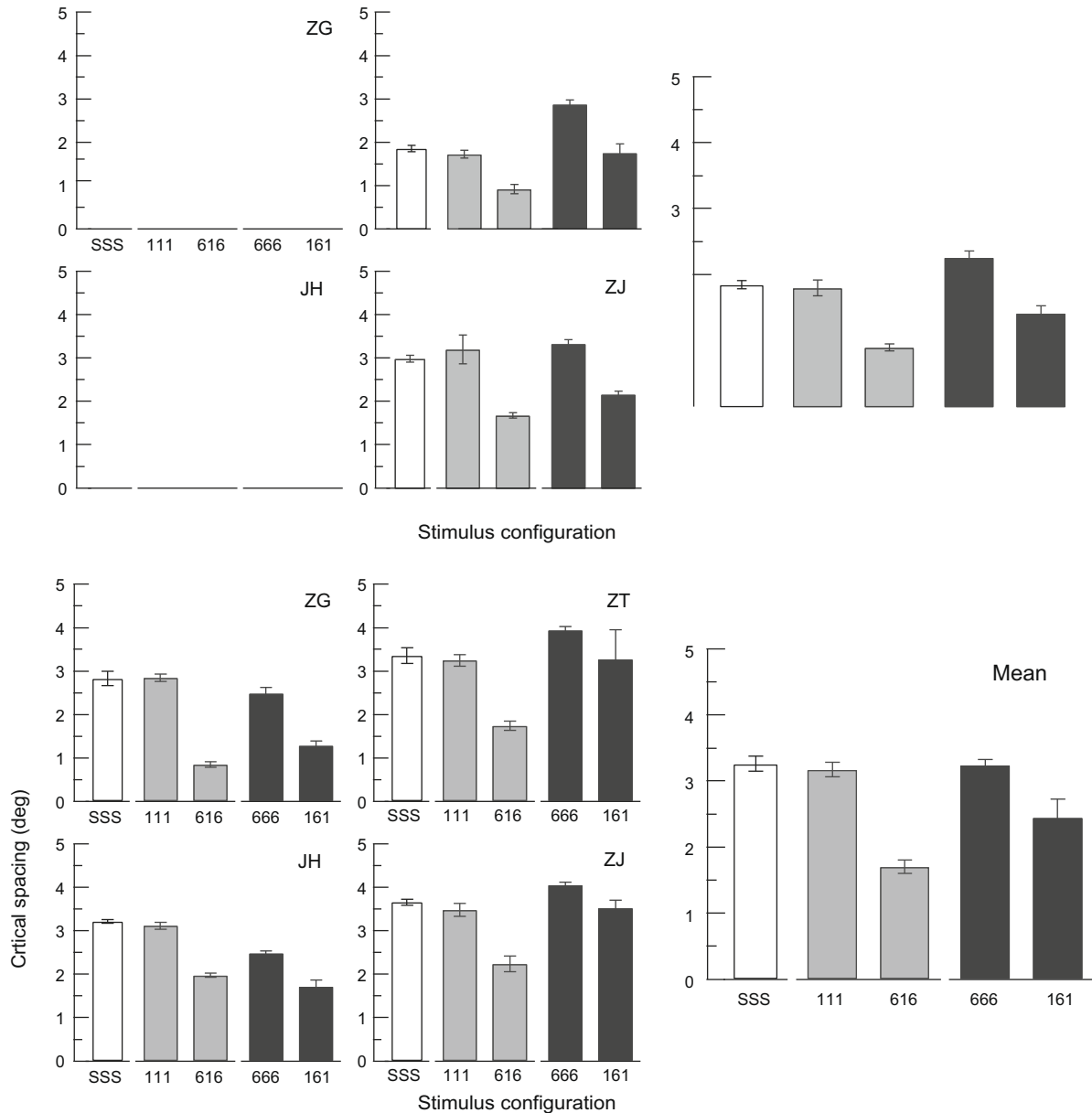
Fig. 6b showed that when the flankers were drawn from a different similar group, crowding was significantly reduced ($F = .007$, repeated measures ANOVA). The mean threshold size was reduced by 27.9 ± 6.3% for CC1 and 19.5 ± 5.6% for Sloan letters. There was no significant difference of crowding reduction between CC and Sloan letter similarity ($F = .221$). These results demonstrated that the observers' knowledge of average and flanker identities as a crowding angle in the center could significantly reduce crowding. However, compared to threshold reduction in the full complex pattern condition (616) size, the complex pattern condition (111), which was 55.5 ± 4.4% (Fig. 4), threshold reduction in the "different flanker condition size, the "same flanker condition at the correction rate of 27.9 ± 6.3% was less robust. This difference suggests that the crowding angle in the center could only account for part of the full complex pattern effect on crowding, and the remaining effect needed to be a robust similarity physical difference that also segregates the average and flankers of the crowding (Ching et al., 2001; Hess et al., 2000; Kooi et al., 1994; Nair, 1992).

Again, the above calculations of thresholds implicitly assumed equal guessing rates of the average in "same" and "different flanker conditions. Under the conditions where both flankers were recognizable, the average guessing rates for the "same" and "different condition could be 1/3 and 1/5, respectively. So the above estimation of the crowding angle in the center of the crowding, which was reduced by the threshold differences between the "same" and "different flanker conditions, was most conservative, as discussed in Experiment III.

3.4.2. The effect of crowding angle on the crowding angle

I have been proposed that crowding results from intermediate-level improper integration of average and flanker features when the average and flankers fall into an integration zone (Lee, Hariharan, & Klein, 2002; Pelli et al., 2004). Having quantified the crowding angle in the center of the crowding, we were able to manipulate the low-level flanker properties to have a close look of this improper feature integration process. Specifically, we measured crowding in the stroke-scrambled CC1 flankers ("stroke condition, Fig. 6), which scrambled the spatial arrangement of the strokes but retained all legible and brush strokes (features), and in the pixel-scrambled CC1 flankers ("pixel condition, Fig. 6), which abolished all legible strokes, and compared threshold changes against the other flanker conditions.

Like the "different flanker conditions, observers could not report the flankers as the average but mistake in the stroke- and pixel-scrambled flanker conditions, so his crowding angle in the center was matched. Moreover, stroke-scrambling broke the low-level processing of flanking characters that would have aided features together, possibly allowing the strokes to be more easily integrated into



he arge . Mean hile, pi el-scrambling des ro ed fea res of he anking charac ers, h s disco raged arge anker fea re in e-gra ion. The res l s sho ed ha s roke-scrambled ankers ("s rkS) raised hreshold si es b 38.4 ± 7.6% compared o hose i h he nscrambled "diff ankers (Fig. 6b; <.001, paired - es), s gges ing ha le er-le el gro ping of anker fea res disco raged arge anker fea re in egration. Moreo er, af er his le er-le el fea re gro ping as disabled b s roke-scrambling of he ankers, he hreshold si es ere no signi can l differen from he "same anker condi ion le el (=.95). I is or h men- ioning ha al hō gh he "same and "s rkS ankers prod ced similar cro ding, cro ding b "s rkS ankers as affect ed b o co n erac ing processes: a op-do n in ence ha red ced cro ding, and a freer arge anker fea re in egration de o disabled le er-le el fea re gro ping ha facili a ed cro ding. S ch d namics ere no discernible i hō a baseline reference of

op-do n impac se b he "diff anker condi ion. On he o her hand, pi el-scrambled ankers ("p lS) nearl iped o cro ding. The hreshold si es ere no signi can l differen from he no- anker baselines (=.086). This effect as predic ed b he fea re in egration model, beca se af er pi el-scrambling, here ere no eligible fea res in he ankers ha co ld be in egrated i h he arge o prod ce cro ding.

4. Discussion

In his s d e demons tra ed i hin-charac er cro ding in recogni ion of isola ed, predominant l comple , CCs in he is al peripher , and sho ed ha s ch i hin-charac er cro ding as rendered negligible b m ch s ronger be een-charac er cro ding once he arge charac er as anked b o her charac ers. We also fo nd red ced cro ding as a res l of spa ial comple i con ras

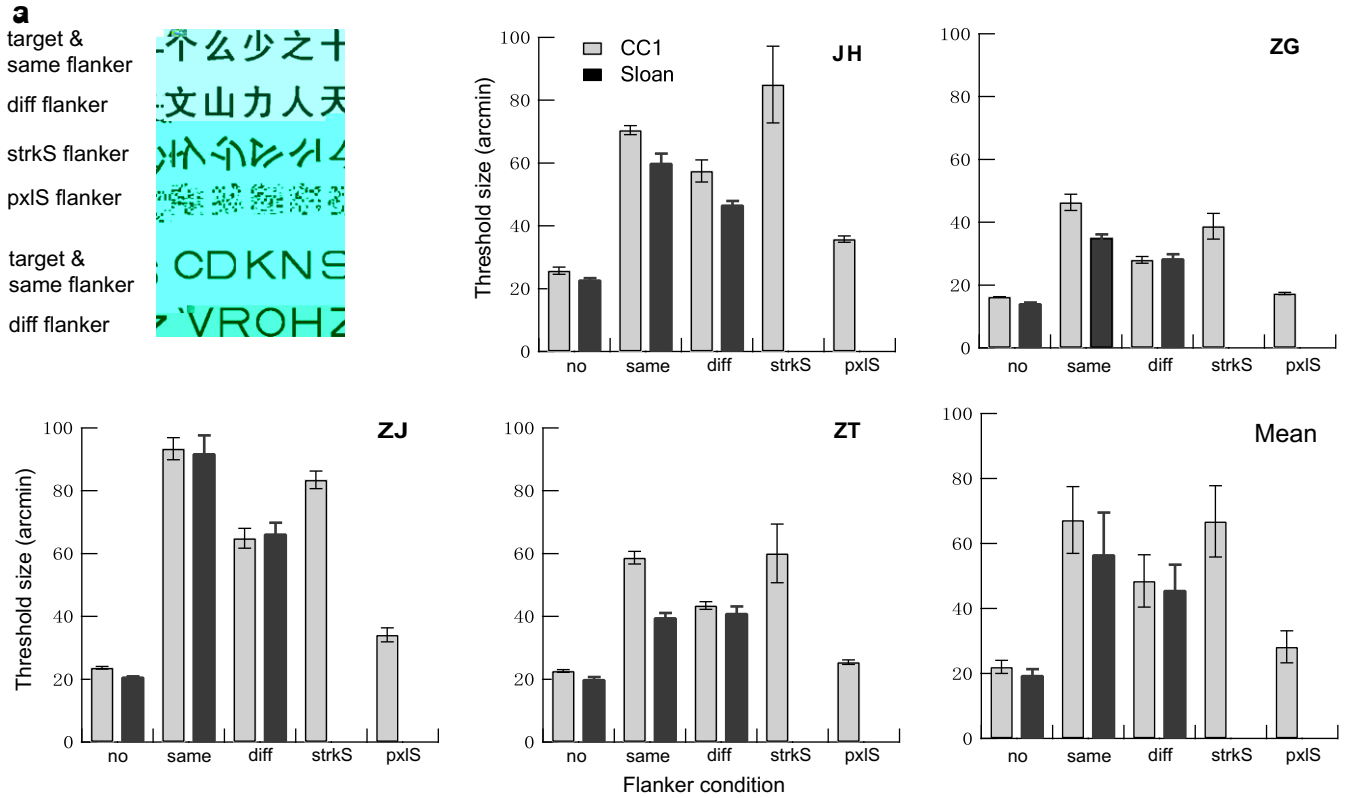


Fig. 6. Top-down and bottom-up influences on crowding. (a) CC1 and Sloan letters used as target and different flanker stimuli. (b) Threshold size as a function of different flanker conditions. no: no-flanker; same: the target and flankers drawn from the same stimulus group; diff: the target and flankers drawn from different stimulus groups; strkS: stroke-scrambled flankers; pxIS: pixel-scrambled flankers.

been the target and flanking CCs, and assessed the contributions of top-down and bottom-up processes to this complex crowding effect and crowding in general.

4.1. We have carried out a series of experiments to...

Our data showed that, as the relative eccentricity increases, complex CCs have to be enlarged at a more rapid rate than simple CCs to reach equal legibility. Complex characters have more strokes than simple ones, and thus have higher object spatial frequency components (cycles/char, Parish & Sperling, 1991). Would the differences in object spatial frequency account for spatial scaling differences among different CC groups?

It is known that visual acuity varies linearly with relative eccentricity (Herse & Bedell, 1989; Levi, Klein, & Aitsebaomo, 1985; Lighthill, 1941; Romano & Virsù, 1979). If S and S_E are cutoff relative frequencies in the fovea and at E degrees eccentricity, then $S_E = S / (1 + E/E_2)$, where E_2 is the eccentricity at which the resolution has changed by a factor of 2. For a character whose height is H degrees and whose object frequency is c/char , its dominant relative spatial frequency is $/Hc/\text{deg}$. When acuity threshold height is reached at an eccentricity E , the character's relative frequency $S_E = /H = S / (1 + E/E_2)$, and the threshold character height should vary with eccentricity in a linear fashion: $H = (1 + E/E_2)/S$. As a result, the acuity height is $H_0 = /S$. If we normalize each character's foveal acuity height H_0 , the normalized acuity height will be $H/H_0 = 1 + E/E_2$, which is independent of the stimulus object frequency, and the normalized lines should all be on top of each other. Thus, the differences in object spatial frequency are not responsible for the steeper scaling of complex CCs in Fig. 2c. Rather, the possibility is that the scaling differences might have resulted

from interactions among pairs of complex CCs, or "dichromatic crowding."

Marelli, Majaj, and Pelli (2005) reported that crowding thresholds for recognition of a feature (a motion or a letter) become higher when the feature is presented in a conjunction (a face or a word) than when it is presented in isolation. This "face and word inferiority effect" appears to occur only in the periphery. Sheed, Sbbaram, Zimmerman, and Hayes (2005) reported a "letter superiority effect," in which high contrast lowercase letters have 10–20% better foveal acuity than words made of 5–6 lowercase letters. In both cases, pairs are more legible when presented alone than when presented in a meaningful whole, which is termed as "intrinsic crowding" by Marelli et al. (2005). Our results revealed a different aspect of the pair-whole relationship, in which a compound object made of more than one meaningful part is more difficult to recognize in the visual periphery than an identifiable simple object. However, further experiments are required to provide direct evidence for crowding in a compound character. Nevertheless, if such interactions exist, they must occur before the whole is recognized. In comparison, the pair-whole superiority effect may occur after the whole is recognized. For this reason, we name the interactions as "dichromatic crowding for discrimination."

Visual character crowding in the periphery may complicate the self-functionality of Chinese reading patterns. In foveal vision, there is a rather simple relationship between the eccentricity and legibility of different complex CCs (Zhang et al., 2007), which allows inference of foveal acuity in recognizing different complex CCs on the basis of one acuity measurement. However, this simple relationship does not apply to the peripheral dichromatic character crowding. A recent series of experiments in China showed that the prevalence of age-related macular degeneration in the

75+ age group is 15–30% (Tian, Zhang, Li, Zhang, & M., 2005). Many of these participants may not all have a reliance on peripheral vision for their daily activities, including reading. Their peripheral visual ability will have to be assessed in a proper consideration of the within-character crowding. On the other hand, in real-world reading materials, CCs are organized in lines with small spacing between them. Our results suggest that within-character crowding may become less important in reading real Chinese because the between-character crowding is likely to dominate (Fig. 3).

4.2. The effect of character eccentricity on crowding

Crowding is markedly reduced when the target and flankers are different in spatial complexity (Fig. 4). Such complexity contrast effects may occur only rarely in the natural alphabets of uniform complexity, but is common in the natural alphabets of Chinese and Japanese. Therefore, the effect of crowding in Chinese may be lower than what is predicted from an experiment using target and flankers of the same complexity.

Boima (1970) showed that when the center letter of a bigram is presented at an eccentricity E , the critical spacing (the center–center spacing between the target and flankers) has produced the same acuity as an isolated letter is roughly $0.5E$. This result has been established in the case of a Latin alphabet, which has a spatial eccentricity dependence on the relative eccentricity of the target. Although the effect of critical spacing is known to depend on the criterion for threshold (Levi, 2008), once a criterion is set, Boima's law would predict similar critical spacing for a given eccentricity regardless of the stimulus types and configurations. We found that the center–center critical spacing varies from $0.23E$ for Sloan letters to $0.37E$ for CC6 characters, the difference of which could be due to within-character crowding in complex CCs. Furthermore, we found that character and critical spacing are significantly reduced in the presence of target–flanker complexity contrast. The changeable critical spacing was also reported by Chang (2007) who demonstrated that critical spacing can be altered through training. These results suggest that relative eccentricity is not the only variable that determines the spatial extent of crowding. Critical spacing may be influenced by multiple factors, and Boima's law, as stated in its original form, may be a special case that holds when stimuli are relatively simple and when the target and flankers share similar spatial complexity.

4.3. The effect of character eccentricity on crowding

Accumulating evidence from many crowding studies including orthographic processing has shown that crowding may be a main cause of visual processing. A number of studies (Levi, 2002) and Pelli et al. (2004) proposed that crowding results from improper integration of target and flanker features in the periphery. The null crowding effect of pixel-scrambled flankers (Fig. 6) is consistent with this account. In addition, the effect of stroke-scrambled flankers (Fig. 6) suggests that target–flanker feature integration is in some measure restricted by the level of processing. Features are set free for integration in the higher-level processing when the higher-level processing is impaired, which aggravates crowding. Previous results (Chang et al., 2001; Hess et al., 2000; Kooi et al., 1994; Naïr, 1992) and orthographic evidence (Fig. 6) also indicated that target–flanker stimulus physical differences help segregate the target and flanker. This stimulus-driven target–flanker segregation likely reduces crowding by restricting the target and flanker features to be in separate channels. This effect is similar to the case in center–center crowding, in which the stimulus-driven and center–center stimuli are grouped into separate channels, center–center crowding is greatly weakened (Malania, Herog, & Westheimer, 2007).

A higher-level processing, or restricted stimulus-driven segregation, may have observers more likely to report a flanker stimulus as the target when a wrong response is made (Stashinsky, 2005). The “same” and “different” flanker effects shown in Fig. 6 indicate that crowding due to this misreporting could be corrected when the observers can separate the target and flanker stimulus through top-down inferences. Stashinsky explained his finding as displaced attention to the flanker location. If this is true, the top-down inference could affect crowding by influencing the position uncertainty of attention. In addition, the same top-down inference could further facilitate target–flanker segregation initially driven by target–flanker physical differences, a possibility we cannot exclude.

A competing explanation of crowding against the improper feature integration model is that crowding could result from limited attentional resolution in the visual periphery (He, Cavanaugh, & Inrillo, 1996; Inrillo & Cavanaugh, 2001). The target becomes less legible when flankers are close because the attentional spotlight is not small enough to separate them. Although these competing models typically make similar predictions about crowding (Levi, 2008), the limited attentional resolution model would have different predictions for the stroke-scrambling effect since the spatial layout of the bigram stimulus is unchanged. However, our evidence is not necessarily against the attentional resolution model since the latter operates at a higher level of visual processing.

Acknowledgments

This research was supported by a National Science Foundation of China Grant (30725018) and a Changjiang Scholar Professorship (C.Y.), and funds from the Beijing Normal University Project 111 (C.Y. & L.L.). We thank Dennis Levi for his helpful comments on this manuscript.

References

- Boima, H. (1970). In-eraction effects in parafoveal letter recognition. *Nature*, 226(5241), 177–178.
- Chang, S. T. (2007). Learning to identify crowded letters: Does improved reading speed? *Vision Research*, 47(25), 3150–3159.
- Chang, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research*, 41(14), 1833–1850.
- Flom, M. C. (1991). Contrast in-eraction and the crowding effect. In R. P. R. S. (Ed.), *Perception* (Vol. 3, pp. 237–257). Philadelphia, PA: JB Lippincott Company.
- Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contrast in-eraction and visual resolution: Contralateral effects. *Science*, 142, 979–980.
- He, S., Cavanaugh, P., & Inrillo, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, 383(6598), 334–337.
- Herse, P. R., & Bedell, H. E. (1989). Contrast sensitivity for letter and grating target under various stimulus conditions. *Optometry and Vision Science*, 66(11), 774–781.
- Hess, R. F., Dakin, S. C., & Kapoor, N. (2000). The foveal ‘crowding’ effect: Psychophysical and physiological? *Vision Research*, 40(4), 365–370.
- Inrillo, J., & Cavanaugh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43(3), 171–216.
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similar and different on spatial in-eraction in peripheral vision. *Stimulus*, 8(2), 255–279.
- Levi, D. M. (2008). Crowding: An essential bottleneck for object recognition: A mini-review. *Vision Research*, 48(5), 635–654.
- Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppression and facilitation of spatial in-eraction in peripheral vision: Peripheral crowding is neither serial nor simple contrast masking. *Journal of Vision*, 2(2), 167–177.
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, 25(7), 963–977.
- Lidghe, E. (1941). Efferent visual acuity as measured with Snellen Test Letters. *Acta Psychologica*, 303–310.
- Malania, M., Herog, M. H., & Westheimer, G. (2007). Grouping of cone elements has affective thresholds. *Journal of Vision*, 7(2), 1–17.
- Marcelli, 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), 58–70.
- Naïr, T. A. (1992). Effects of lateral masking and spatial precision on gap-resolution in central and peripheral vision. *Vision Research*, 32(4), 771–777.

- Parish, D. H., & Sperling, G. (1991). Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Research*, 31(7–8), 1399–1415.
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Dissociating feature integration from detection. *Journal of Vision*, 4(12), 1136–1169.
- Roman, J., & Virsù, V. (1979). An estimation and application of the human cortical magnification factor. *Electroencephalography and Clinical Neurophysiology*, 37(3), 495–510.
- Sheedy, J. E., Sbbaram, M. V., Zimmerman, A. B., & Hayes, J. R. (2005). Text legibility and the letter superiority effect. *Human Factors*, 47(4), 797–815.
- Straussburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect foveation. *Journal of Vision*, 5(11), 1024–1037.
- Strauss, J. A., & Brian, H. M. (1962). A study of separation difficulty: Its relationship to visual acuity in normal and amblyopic eyes. *Acta Ophthalmologica*, 53, 471–477.
- Tian, M., Zhang, Y., Li, L., Zhang, C., & Meng, Y. (2005). An epidemiological investigation of age-related macular degeneration. *Journal of the Chinese Medical Association*, 31(2), 70–71.
- Tripathy, S. P., & Cananagh, P. (2002). The effect of crowding in peripheral vision does not scale with retinal eccentricity. *Vision Research*, 42(20), 2357–2369.
- Wolford, G., & Hollingworth, S. (1974). Retinal location and starting position as important variables in visual information processing. *Perception & Psychophysics*, 16, 437–442.
- Zhang, J. Y., Zhang, T., Xie, F., Li, L., & Yu, C. (2007). Legibility variations of Chinese characters and implications for visual acuity measurement in Chinese reading population. *Journal of the Chinese Medical Association*, 31(5), 2383–2390.