Chinese-character crowding—I. Effects of structural similarity

Leo Yuk Ting Cheung Department of Psychology, The University of Hong Kong, Hong Kong

Sing-Hang Cheung Department of Psychology, The University of Hong Kong, Hong Kong

Crowding impedes the identification of flanked objects in peripheral vision. Prior studies have shown crowding strength decreases with target–flanker similarity. Research on crowding in Chinese-character recognition has been scarce in the literature. We aimed to fill the research gap by examining the effects of structural similarity on Chinese-character crowding. Regularity in within-character configuration, i.e., *orthographic legality*, of flankers was manipulated in Experiment 1. Target–flanker similarity in orthographic legality did not affect crowding strength, measured as contrast threshold elevation. Crowding weakened only when the strokes in the flankers were scrambled. Contour integrity of flankers was manipulated by randomly perturbing the phase spectra of the stimulus images in Experiments 2a and 2b. Crowding by perturbed-phase flankers remained robust but was weaker compared with intact-phase flankers. Target–flanker similarity in contour integrity modulated crowding strength. Our findings were consistent with the postulation that faulty integration of low-level visual features contributed to crowding of Chinese characters. Studies on Chinese-character recognition and crowding can provide important insights into how the visual system processes complex daily objects.

Introduction

When objects are placed close to one another in peripheral vision, they can still be detected but seem cluttered due to visual crowding (Bouma, 1970; Stuart & Burian, 1962; see also Levi, 2008; Whitney & Levi, 2011, for reviews). Although the exact mechanism behind crowding is still under debate, many of its properties have been well studied. In the current study, we studied one such property—similarity effects—in crowding of Chinese characters as we manipulated the structural properties of flankers.

According to the faulty feature integration theory of crowding, flanker features that fall into the *integration field* of a target will be erroneously combined with the target, especially in parafoveal vision (Levi, Hariharan, & Klein, 2002; Nandy & Tjan, 2007; Pelli, Palomares, & Majaj, 2004). The size of the integration field is known to increase with target eccentricity (Toet & Levi, 1992). Such observation has led researchers to link the theoretical concept of the integration field with a potential physiological counterpart, the receptive field of certain visual neurons (Freeman & Simoncelli, 2011; van den Berg, Roerdink, & Cornelissen, 2010). Computational modeling showed that faulty integration through compulsory averaging between the target and flanker signals could explain crowding of carefully designed stimuli carrying orientation and positional information (Greenwood, Bex, & Dakin, 2009; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001).

Crowding strength is known to increase with target–flanker similarity: Crowding becomes stronger when the target and flankers are similar in terms of shape (Kooi, Toet, Tripathy, & Levi, 1994; Nazir, 1992), orientation (Andriessen & Bouma, 1976; Levi et al., 2002), spatial frequency (Chung, Levi, & Legge, 2001), contrast polarity (Chung & Mansfield, 2009; Kooi et al., 1994), color (Kooi et al., 1994; Põder, 2007), or visual complexity (Zhang, Zhang, Xue, Liu, & Yu, 2009). Bernard and Chung (2011) found that a target letter was more susceptible to crowding when flanked by more confusable letters. Applying the faulty feature integration theory, features of flankers that were visually similar to the target could be more easily integrated with those in the target.

Studies on crowding of Chinese characters were relatively scarce in the literature. Chinese is the most widely used logographic language, and its characters have far more varied spatial complexity than letters in an alphabetic language. Among the 5,000 most commonly used Chinese characters in Hong Kong (Hong Kong SAR Education Bureau, 2012), the number of strokes ranges from one to 32. A major psychophysical implication of such variation is that Chinese characters with high complexity require a

Citation: Cheung, L. Y. T., & Cheung, S.-H. (2017). Chinese-character crowding—I. Effects of structural similarity. *Journal of Vision*, *17*(11):14, 1–13, doi:10.1167/17.11.14.

doi: 10.1167/17.11.14 Received May 13, 2017; published September 29, 2017 ISSN 1534-7362

larger size to reach the same level of acuity in both central and peripheral vision (Zhang, Zhang, Xue, Liu, & Yu, 2007). More importantly, consistent with the principle of similarity effect, flankers with a spatial complexity level different from the target were found to produce weaker crowding effect (Zhang et al., 2009).

In addition to having a wide range of spatial complexity, Chinese characters are in a special class of objects with a well-defined hierarchy of orthographic components. On the most basic level, a Chinese character could be decomposed into strokes. Some of these strokes combine to form stroke patterns: Some are radicals that signify semantic classification of the character; others might be phonetic representations. The legality of a Chinese character is determined by both the validity of the individual components and their spatial arrangement within an imaginary square frame. In Chinese-character recognition, observers not only process the components of Chinese characters, but also the spatial relationship among these components although the former could be more important to Chinese readers (Hsiao & Cottrell, 2009). In the current study, we studied the role of object-level information in terms of orthographic configuration in crowding of Chinese characters.

Object-level information, in addition to feature-level information, has been shown to influence crowding strength in face crowding (Louie, Bressler, & Whitney, 2007). Like Chinese characters, a face is also defined by its components and their configuration. Louie et al. (2007) found that crowding of an upright face was stronger with upright face flankers than with inverted face flankers. On the other hand, inverted face flankers did not result in stronger crowding of an inverted face target. Therefore, Louie et al. concluded that this was not merely a similarity effect. Instead, object-level configural information in upright faces played a role in face crowding. Here we examined whether similarity in orthographic legality, a form of object-level information, would affect the strength of Chinese-character crowding. We compared the crowding strengths as we manipulated orthographic legality in Experiment 1. We then followed up on what we found in Experiment 1 and studied the effect of similarity in contour integrity in Experiments 2a and 2b.

Experiment 1: Similarity in orthographic legality

Zhang et al. (2009) found that crowding was inversely proportional to orthographic scrambling: Pixel-scrambled character flankers, devoid of any orthographic information, induced almost no crowding, and stroke-scrambled character flankers resulted in weaker crowding than intact characters. The current experiment further investigated if the legality of stroke pattern arrangement would affect crowding. We compared crowding strength from legal character flankers versus flanking characters with illegal configuration of valid stroke patterns. If similarity effects on crowding could be extended to such high-level factors as orthographic legality, flankers of higher configurational legality (i.e., more similar to the target) would induce stronger crowding. Alternatively, configurational legality of flankers might have no impact on crowding strength if crowding mostly occurred at low-level visual processing.

Methods

Observers

Ten native Chinese readers, aged between 18 and 33, participated in Experiment 1. All observers had normal or corrected-to-normal vision. Time spent in this experiment was compensated for HK$60 an hour. Written informed consent was collected from each observer before the experiment. All procedures used in this experiment were reviewed and approved by the Human Research Ethics Committee of the University of Hong Kong and adhered to the tenets of the Declaration of Helsinki.

Stimuli

Chinese characters were selected from a list of 3,000 most frequently used traditional Chinese characters in Hong Kong (Poon & Kang, 2003). These frequently used characters were further screened for the configuration of their constituent stroke patterns. A group of 176 "left–right" characters that were composed of a left and a right stroke pattern were included in the final set of *real* characters. Among all left–right Chinese characters, some stroke patterns can exist only on the left side of a character, and others can exist only on the right side. Forty-five *pseudo*characters were created by combining the left-only and right-only stroke patterns in a legitimate way (i.e., the left-only and right-only stroke patterns stayed in the corresponding side of the character), yet such combinations were not existing characters. Forty-five *non*characters were created by swapping the left and right stroke patterns in each of the 45 pseudocharacters. The real-, pseudo-, and noncharacters shared 28 different left stroke patterns and 45 right stroke patterns. Total stroke number ranged from four to 13 ($M = 8.70$, $SD = 1.89$) for real characters and from six to 12 ($M = 8.60$, $SD = 1.63$) for both pseudo- and noncharacters. In the Chinese writing system, a stroke is a continuous movement of the writing instrument with a disjunction in the end, which

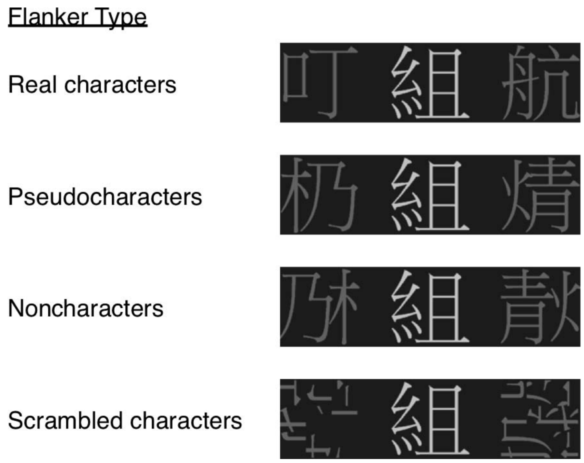

Figure 1. Examples of stimuli used in the four flanked conditions of Experiment 1. The target (middle character) was always a real Chinese character. The flankers varied in levels of orthographic legality, going from real characters, pseudocharacters, and noncharacters to scrambled characters.

allows corners and hooks. In addition to stroke count, visual complexity of the stimuli was also measured through perimetric complexity, which is the ratio of the squared sum of inside and outside perimeters to the ink area of the character (Attneave & Arnoult, 1956; Pelli, Burns, Farell, & Moore-Page, 2006). Mean perimetric complexities were 465.85 ($SD = 64.28$) for real characters, 459.60 ($SD = 60.46$) for pseudocharacters, and 455.89 ($SD = 63.76$) for noncharacters. A set of 176 *scrambled* characters were created by dividing a real character into nine square units of equal size and having these units randomly rotated and swapped. As our scrambling method could break up strokes into smaller fragments, mean perimetric complexity among scrambled characters was 548.06 ($SD = 82.97$). Figure 1 shows examples of the four types of characters. The characters were rendered and presented in the PMingLiU font (*Xin Si Ming Ti*). Height and width of each character were both 1.5°, which exceeded the size threshold of all observers ($M = 0.75°$, $SD = 0.19°$).

Apparatus

This experiment was programmed and run in the MATLAB environment (R2008b, MathWorks, MA) with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli were presented on a 17-in. Dell M783c CRT monitor (1024 × 768 pixels, 85-Hz frame rate) controlled by a Mac Pro computer. The output luminance of the monitor was calibrated and linearized. Background luminance was 13.5 cd/m^2. A chin rest was used to maintain a viewing distance of 100 cm.

Procedures

Each observer performed three blocks of 40 trials for each of the five conditions—four flanked conditions and a target-only baseline condition. The target in all conditions was always a real character. Flankers were presented to the left and right of the target with a center-to-center distance of 2° in all flanked conditions. Both flankers were real, pseudo-, non-, or scrambled characters depending on the stimulus condition. The order of conditions was randomized across the observers.

Contrast threshold for 75% accuracy in target recognition was estimated in each block using QUEST, a Bayesian adaptive procedure (Watson & Pelli, 1983). The threshold criterion was selected such that performance would be well above chance and comparable to criteria used in previous studies. Michelson contrast of the target remained at 0.30 for the first five trials in each block. Target contrast in each trial was then determined by QUEST for the remaining 35 trials. Flanker contrast was fixed at 0.10. Flanker contrast level was chosen based on our pilot data on contrast threshold in the target-only baseline condition.

Observers were instructed to fixate at a cross on the screen at the beginning of each trial and to initiate the trial by a key press. After a trial was initiated, the fixation cross continued to remain on the screen for 200 ms, followed by a blank screen for 300 ms before the stimulus was presented for 100 ms. The target character was always presented at 6° eccentricity in the right visual field. A response screen with the correct response and 22 randomly drawn characters was shown at the end of each trial. The observers used the arrow keys on the keyboard to choose a response from the set of 23 characters. Auditory feedback was given after each trial. Figure 2 illustrates the sequence of a trial.

Statistical analysis

We analyzed the data using R (Version 3.3.0; R Core Team, 2016). Linear mixed effects models with random effects of the observers were fitted and examined using the nlme (Version 3.1-128; Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2016) package. Significance level in planned pairwise comparisons was adjusted through deriving a multivariate normal distribution for the null hypotheses to control the family-wise error rate (Hothorn, Bretz, & Westfall, 2008). For effect size measures, marginal R^2 and conditional R^2, which estimated the proportion of variances accounted for by the fixed effects only and the full model (i.e., fixed effects plus random effects), respectively, were computed using the piecewiseSEM package (Lefcheck, 2015). Hedge's g, computed using the effsize package (Torchiano, 2016), was reported alongside with each pairwise comparison. Confidence intervals around the

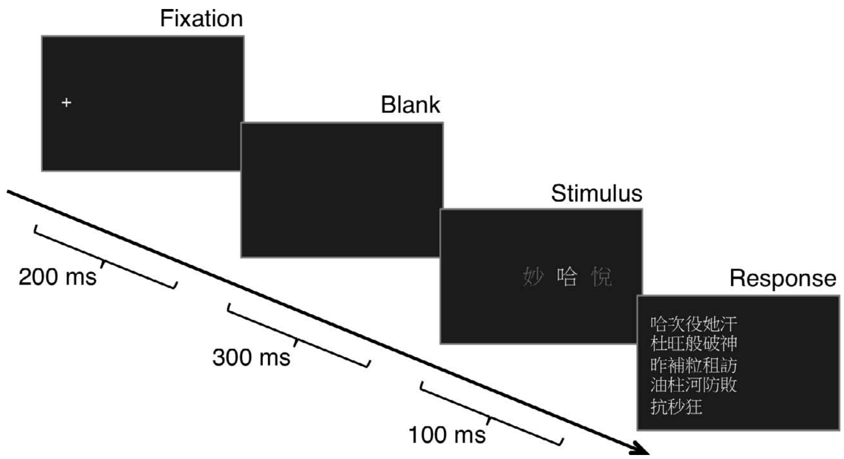

Figure 2. Illustration of the Chinese-character identification task used in this study. The stimulus was either a single character or a trigram of characters presented on the horizontal meridian in the right visual field. Flanking characters of different types were used in different experiments. A response screen showing 23 (Experiment 1) or 20 (Experiments 2a and 2b) choices was displayed at the end of the trial sequence until a response was made. The characters shown were used in Experiment 1 only and were less complex than those used in Experiments 2a and 2b (see Figure 4 for examples).

means were constructed through bootstrapping using the boot package (Version 1.3-18; Canty & Ripley, 2016). Bias-corrected and accelerated (BC_a) bootstrap confidence intervals (Efron, 1987) were reported.

Results

Mean contrast threshold for the target-only baseline condition was 0.08 [$SD = 0.03$, 95% BC_a CI = (0.06, 0.10), range = 0.03–0.14]. Crowding strength was measured by the elevation of contrast threshold when we compared the flanked conditions against the target-only condition. Figure 3 shows the results of Experiment 1. Mean threshold elevations of the real-, pseudo-, non-, and scrambled-character flanked conditions were 3.59 [$SD = 1.46$, 95% BC_a CI = (2.62, 4.35)], 3.08 [$SD = 1.42$, 95% BC_a CI = (2.44, 4.19)], 3.47 [$SD = 1.39$, 95% BC_a CI = (2.61, 4.23)] and 2.68 [$SD = 1.14$, 95% BC_a CI = (2.08, 3.41)] respectively. All threshold elevations exceeded one, indicating robust crowding in all flanked conditions.

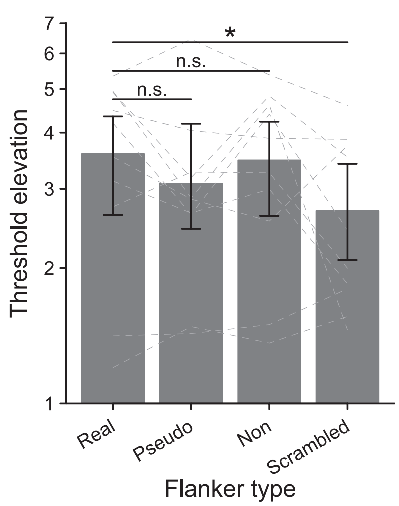

Figure 3. Mean threshold elevation in different flanked conditions of Experiment 1 ($N = 10$). Error bars represent 95% bias-corrected and accelerated (BC_a) bootstrap confidence intervals. Data from individual observers are also plotted as dashed lines. *Two-tailed $p < 0.05$.

A linear mixed effects model with the fixed effect of flanker types and random effect of observers was fitted to the log (base 10) threshold elevations (marginal $R^2 = 0.03$; conditional $R^2 = 0.53$). Shapiro–Wilk test suggested no statistically significant deviation from normality in the residuals, $W = 0.987$, $p = 0.325$. Homoscedasticity assumption was assessed by comparing models fitted with and without equal within-condition variance. The heteroscedastic model was not significantly better, $\chi^2(1) = 3.12$, $p = 0.374$. The omnibus test found no statistically significant difference among the mean threshold elevations across the four flanked conditions, $F(3, 107) = 2.22$, $p = 0.090$.

Planned comparisons were also conducted comparing the pseudo-, non-, and scrambled-character flanked conditions against the real-character flanked condition. Compared with the real-character flanked condition, scrambled-character flanked condition had significantly lower threshold elevation [mean difference = −0.91, $SD = 1.18$, 95% BC_a CI = (−1.76, −0.35), $z = -2.41$, $p = 0.042$, Hedge's $g = -0.74$], but the pseudocharacter flanked condition [mean difference = −0.50, $SD = 1.04$, 95% BC_a CI = (−1.14, 0.09), $z = -1.32$, $p = 0.407$, Hedge's $g = -0.46$], and noncharacter flanked condition [mean difference = −0.11, $SD = 0.43$, 95% BC_a CI =

(−0.42, 0.11), $z = -0.51$, $p = 0.923$, Hedge's $g = -0.25$], did not.

Discussion

Crowding remained strong in the pseudocharacter and noncharacter conditions but weakened when the stroke patterns were fragmented in the scrambled-character condition. Our results were consistent with Zhang et al.'s (2009) study, in which stroke-scrambled flankers resulted in weaker crowding. Strength of Chinese-character crowding does not depend on the orthographic legality of flankers. As long as the high-level form of stroke patterns remains intact, flankers induce strong crowding.

Contour alignment was found to modulate crowding strength (Chakravarthi & Pelli, 2011; Glen & Dakin, 2013) and might contribute to the weakened crowding in the scrambled-character crowded condition. Rotation was involved when we scrambled the characters and could have made alignment of strokes from target and from flankers less likely. Greenwood, Bex, and Dakin (2012) also reported that crowding of stroke orientation judgment was strongest when orientation of the flanker strokes was most similar to that of the target stroke. Changes in the orientation distribution of strokes in the scrambled flankers might partially explain the observed weaker crowding.

When the form of a visual object is distorted, the distribution of contrast energy across different spatial frequencies may also be disrupted. Would crowding remain strong if only the visual form of the flankers is distorted but with an intact contrast energy distribution? We addressed this in our next experiment.

Experiment 2a: Similarity in contour integrity

Chinese characters can be decomposed not only into different orthographic components, but also into low-level features of different spatial frequencies and orientations. Using Fourier transform, any image can be decomposed into its phase and magnitude spectra. Between the two, the phase spectrum is more important to retaining features in an image (Oppenheim & Lim, 1981; Wang & Simoncelli, 2004). Nevertheless, no difference in crowding strength was observed between flankers with an intact and a scrambled phase spectra (C. He & Tjan, 2004; Shin, Wallace, & Tjan, 2010). Tjan and colleagues (C. He & Tjan, 2004; Shin et al., 2010) reported that significant reduction in crowding strength was only observed when the magnitude spectrum was altered. Tjan and colleagues' findings were consistent with the primary visual cortex (V1) being the cortical locus of crowding because V1 does not differentiate between intact-phase and perturbed-phase stimuli (Ben-Shachar, Dougherty, Deutsch, & Wandell, 2007). In Experiments 2a and 2b, we examined the effects of contour integrity through phase perturbation on crowding strength in Chinese characters.

Methods

Observers

Five native Chinese readers, aged between 18 and 33, participated in Experiment 2a. All observers had normal or corrected-to-normal vision. Time spent in this experiment was compensated for HK$60 an hour. Written informed consent was collected from each observer before the experiment. All procedures used in this experiment were reviewed and approved by the Human Research Ethics Committee of the University of Hong Kong and adhered to the tenets of the Declaration of Helsinki.

Stimuli

One hundred ninety high-complexity Chinese characters were selected from a list of 3,171 most frequently used traditional Chinese characters for primary education in Hong Kong (Hong Kong SAR Education Bureau, 2009). A total of 190 high-complexity characters rendered in PMingLiU font (*Xin Si Ming Ti*) were selected for this experiment. Mean stroke count for this group of characters was 17.67 ($SD = 0.81$), and mean perimetric complexity was 651.52 ($SD = 44.91$). These characters were randomly used as targets, flankers, and response options.

We manipulated the phase and magnitude spectra of flanker characters by performing discrete Fourier transform. Random numbers drawn from the von Mises distribution (i.e., the circular normal distribution) with location parameter of zero and dispersion parameter of two were added to the phase spectrum of each character. Each perturbed phase spectrum was combined with the original magnitude spectrum of the corresponding character to create the perturbed-phase characters through inverse Fourier transform. Phase perturbation distributed contrast energy over the whole square area and made the perturbed-phase characters perceptually distinct from the intact-phase characters. We introduced conditions in which the original phase spectrum and the perturbed phase spectrum were combined with the average of the 190 magnitude spectra. With the averaged magnitude spectrum, some of the contrast energy spread over to the background even without phase perturbation. Therefore, with

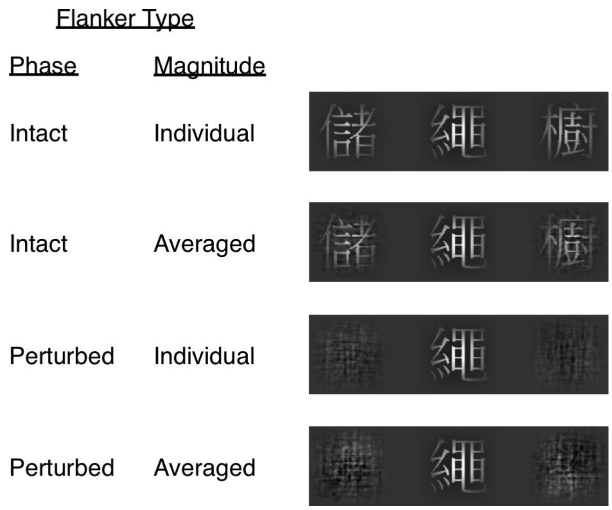

Figure 4. Examples of stimuli used in the four flanked conditions of Experiment 2a. The target (middle character) was always an intact character. The flankers varied in status of phase and magnitude spectra after undergoing Fourier transform: The phase spectrum was either intact or perturbed, and the magnitude spectrum was taken from individual character (the actual flanking character) or averaged from the entire character set.

manipulations on both the phase (intact vs. perturbed) and magnitude (individual vs. averaged) spectra, four types of flankers were created: intact phase with individual magnitude (i.e., without Fourier transform), intact phase with averaged magnitude, perturbed phase with individual magnitude, and perturbed phase with averaged magnitude (see Figure 4 for examples). The image of each character was filtered by a raised cosine function to remove the sharp edges. Height and width of each character image, including small zero padding at the edges, were both 2°.

Apparatus

The same apparatus setup was used as in Experiment 1 except that background luminance was 19.5 cd/m^2.

Procedures

Each observer performed two blocks of flanked trials and two blocks of target-only trials. Trials for the four flanked conditions were randomly intermixed in a block with 60 trials per condition. Sixty trials were performed separately in each target-only block. The target in all conditions was always a character with intact phase and its original magnitude spectrum and presented at 6° eccentricity in the right visual field. Flankers were presented to the left and right of the target with a center-to-center distance of 2° in all flanked conditions.

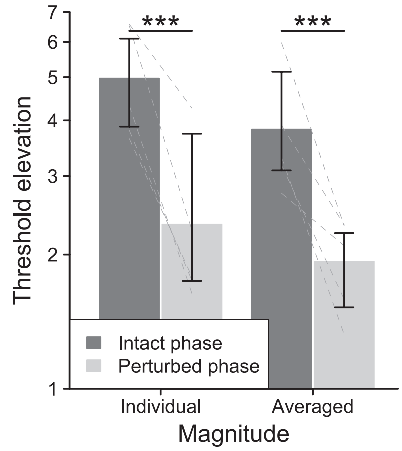

Figure 5. Mean threshold elevation in different flanked conditions of Experiment 2a ($N = 5$). Error bars represent 95% bias-corrected and accelerated (BC_a) bootstrap confidence intervals. Data from individual observers are also plotted as dashed lines. ***Two-tailed $p < 0.001$.

RMS contrast of the target was 0.10, 0.20, or 0.30 for the first 30 trials of each condition in each block. Target contrast in each trial was then determined by QUEST for the remaining trials and used to estimate the contrast threshold for 75% recognition accuracy (Watson & Pelli, 1983). Flanker RMS contrast was fixed at 0.20. No feedback was given in Experiment 2a. Stimulus presentation sequence in each trial was the same as in Experiment 1 except that the response screen presents 20 alternatives, including the target (in both target-only and flanked conditions) and the flankers (only in flanked conditions).

Statistical analysis

The methods of statistical analysis were identical to those in Experiment 1.

Results

Mean contrast threshold for the target-only baseline condition was 0.06 [$SD = 0.03$, 95% BC_a CI = (0.05, 0.10), range = 0.05–0.11]. Crowding strength was measured by the elevation of contrast threshold when we compared the flanked conditions against the target-only condition. Figure 5 shows the results of Experiment 2a. Mean threshold elevations of the intact-phase, individual-magnitude; perturbed-phase, individual-magnitude; intact-phase, averaged-magnitude; and perturbed-phase, averaged-magnitude flanked conditions were 4.97 [$SD = 1.46$, 95% BC_a CI = (3.87, 6.10)], 2.34 [$SD = 1.10$, 95% BC_a CI = (1.75, 3.74)], 3.82 [$SD = 1.27$, 95% BC_a CI = (3.09, 5.14)] and 1.93 [$SD = 0.45$, 95% BC_a CI = (1.52, 2.23)], respectively. All threshold elevations exceeded one, indicating robust crowding in all flanked conditions.

A linear mixed effects model with the fixed effects of flanker phase perturbation and flanker magnitude averaging and random effect of observers was fitted to the log (base 10) threshold elevations (marginal $R^2 = 0.48$; conditional $R^2 = 0.62$. Shapiro–Wilk test suggested no statistically significant deviation from normality in the residuals, $W = 0.949$, $p = 0.070$. Homoscedasticity assumption was assessed by comparing models fitted with and without equal within-condition variance. The heteroscedastic model was not significantly better, $\chi^2(3) = 6.38$, $p = 0.094$. The omnibus test found a statistically significant main effect of flanker phase perturbation, $F(1, 32) = 46.04$, $p < 0.001$, but not of flanker magnitude averaging, $F(1, 32) = 2.68$, $p = 0.111$, on threshold elevation. The interaction between phase perturbation and magnitude averaging was not statistically significant, $F(1, 32) = 0.62$, $p = 0.436$.

Planned comparisons were also conducted comparing the flanked intact-phase and flanked perturbed-phase conditions within each of the two magnitude-averaging conditions. When individual magnitude was used, the flanked perturbed-phase condition had significantly lower threshold elevation than the intact-phase flanked condition [mean difference $= -2.63$, $SD = 0.96$, 95% BC_a CI = (–3.81, –2.12), $z = -5.36$, $p < 0.001$, Hedge's $g = -2.48$]. A statistically significant difference between the perturbed-phase and the intact-phase flanked conditions was also found when averaged magnitude was used [mean difference $= -1.89$, $SD = 1.10$, 95% BC_a CI = (–2.99, –1.22), $z = -4.24$, $p < 0.001$, Hedge's $g = -1.55$].

Discussion

Crowding was weakened with perturbed-phase flankers. Our results suggest that the strength of crowding in Chinese characters is contingent on similarity in terms of contour integrity. Whereas prior studies (C. He & Tjan, 2004; Shin et al., 2010) observed no difference in crowding strength between intact-phase and perturbed-phase flankers, we found otherwise: Perturbed-phase flankers induced weaker crowding than flankers with intact phase even after controlling for difference in magnitude spectrum among the flankers in the averaged magnitude conditions. Contour integrity therefore plays a role in determining crowding strength in Chinese characters.

In previous studies in which null effect on phase perturbation was found, crowding strength was weakened only when the amplitude spectrum or power spectral density was altered. Our findings, however, suggest that contour integrity (phase) also influences the similarity effects in crowding of Chinese characters. The discrepancy in findings might lie in stimulus type. Our stimuli were more complex than average Chinese characters. Like many objects of high visual complexity, the components (i.e., radicals and stroke patterns) are essential to recognition, but the configuration of individual components (cf. faces) is just as important. Thus, Chinese-character recognition requires relatively high precision in feature binding. One possibility is that, with more features in the target, the likelihood of erroneously binding with flanker features before reaching object-level visual processing increases.

Neuroimaging evidence supports the notion that crowding occurs at multiple stages (Anderson, Dakin, Schwarzkopf, Rees, & Greenwood, 2012; Freeman, Donner, & Heeger, 2011). For similarity in a visual element to influence crowding, the element needs to have a crucial role in the process of feature integration. From the discrepancy in findings on phase integrity, we infer that phase may be more involved in feature integration of one type of object but not another. In other words, the features that are erroneously integrated in the crowding process may also differ. This difference can be due to the characteristics of the object and even the requirements of the task. Unlike Latin letters, strokes and stroke patterns can be combined to form numerous different Chinese characters. Phase perturbation reduces the visibility of strokes in Chinese characters and may *reduce* the number of intact strokes to be erroneously integrated in crowding.

Experiment 2b: Effect of trial blocking

Phase perturbation of flankers was found to alleviate crowding of Chinese characters in Experiment 2a. Such findings differed from Tjan and colleagues' findings on crowding of Latin letters (C. He & Tjan, 2004; Shin et al., 2010). In addition to the different stimulus types used, trials of all the flanked conditions were randomly intermixed in a block in Experiment 2a but were performed in separate blocks in Tjan and colleagues' studies. We did a follow-up experiment to examine whether trial blocking contributed to the discrepancy in findings.

In this experiment, we examined threshold elevations brought by intact-phase and perturbed-phase flankers measured in separate blocks of trials of different conditions. When trials of different flanked conditions are performed in separate blocks, the visual system may be able to tune the recognition process according to the type of flankers used in a block. Such flanker-specific tuning of the target recognition process may complicate the comparison between different flanked conditions.

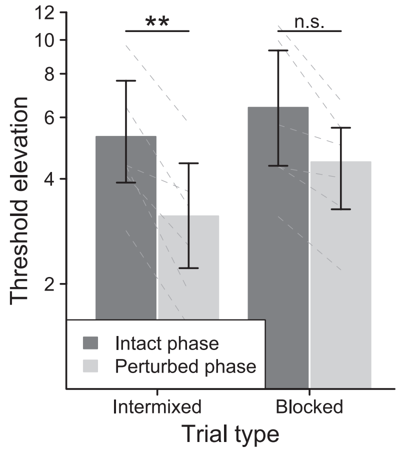

Figure 6. Mean threshold elevation in different flanked conditions of Experiment 2b ($N = 6$). Error bars represent 95% bias-corrected and accelerated (BC_a) bootstrap confidence intervals. Data from individual observers are also plotted as dashed lines. **Two-tailed $p < 0.01$.

Here we asked if trial blocking could explain the null effect of phase perturbation in Tjan and colleagues' (C. He & Tjan, 2004; Shin et al., 2010) studies.

The intact-phase and perturbed-phase conditions in Experiment 2a also differed in the possible roles of mislocation errors. Flankers can be misreported as the target in crowding. Because perturbed-phase flankers could not be recognized, misidentification of flankers as the target was not possible. As a result, crowding with perturbed-phase flankers could be weaker than crowding with intact-phase flankers. Flankers were excluded from the response set in this follow-up experiment to eliminate mislocation errors.

Methods

Observers

Six native Chinese readers, aged between 18 and 33, participated in Experiment 2b. All observers had normal or corrected-to-normal vision. Time spent in this experiment was compensated for HK$60 an hour. Written informed consent was collected from each observer before the experiment. All procedures used in this experiment were reviewed and approved by the Human Research Ethics Committee of the University of Hong Kong and adhered to the tenets of the Declaration of Helsinki.

Stimuli and apparatus

The apparatus setup and stimuli were the same as those in Experiment 2a. Only stimuli with individual magnitude spectrum were used.

Procedures

In addition to the target-only baseline condition, only intact-phase, individual-magnitude and perturbed-phase, individual-magnitude flanked conditions were included. The intact-phase and perturbed-phase trials were either randomly intermixed in a block or were performed in separate blocks. The target-only trials were always performed in a separate block. Sequence of blocks was randomized across observers. The procedures in Experiment 2b were otherwise identical to those in Experiment 2a.

Statistical analysis

The methods of statistical analysis were identical to those in Experiment 1.

Results

Mean contrast threshold for the target-only baseline condition was 0.04, $SD = 0.02$, 95% BC_a CI = (0.03, 0.06), range = 0.03–0.07. Crowding strength was measured by the elevation of contrast threshold when we compared the flanked conditions against the target-only condition. Figure 6 shows the results of Experiment 2b. Mean threshold elevations of the intact-phase-intermixed, perturbed-phase-intermixed, intact-phase-blocked, and perturbed-phase-blocked flanked conditions were 5.29 [$SD = 2.42$, 95% BC_a CI = (3.90, 7.64)], 3.13 [$SD = 1.53$, 95% BC_a CI = (2.22, 4.43)], 6.41 [$SD = 3.27$, 95% BC_a CI = (4.36, 9.33)] and 4.47 [$SD = 1.63$, 95% BC_a CI = (3.27, 5.60)], respectively. All threshold elevations exceeded one, indicating robust crowding in all flanked conditions.

A linear mixed effects model with the fixed effects of flanker phase perturbation and trial blocking and random effect of observers was fitted to the log (base 10) threshold elevations (marginal $R^2 = 0.19$; conditional $R^2 = 0.52$). Shapiro–Wilk test suggested no statistically significant deviation from normality in the residuals, $W = 0.977$, $p = 0.466$. Homoscedasticity assumption was assessed by comparing models fitted with and without equal within-condition variance. The heteroscedastic model was not significantly better, $\chi^2(3) = 0.63$, $p = 0.889$. The omnibus test found statistically significant main effects of flanker phase perturbation, $F(1, 39) = 12.18$, $p = 0.001$, and of trial blocking, $F(1, 39) = 5.70$, $p = 0.022$, on threshold elevation. The interaction between phase perturbation and trial blocking was not statistically significant, $F(1, 39) = 0.64$, $p = 0.428$.

Planned comparisons were also conducted comparing the flanked intact-phase and flanked perturbed-phase conditions within each of the two trial blocking conditions. When flanked trials were randomly intermixed in a block, the flanked perturbed-phase condition had significantly lower threshold elevation than the intact-phase flanked condition [mean difference = –2.16, $SD = 1.19$, 95% BC_a CI = (–3.11, –1.37), $z = -3.03$, $p = 0.005$, Hedge's $g = -1.67$]. A statistically significant difference between the perturbed-phase and the intact-phase flanked conditions was not found when flanked trials were performed in separate blocks [mean difference = –1.93, $SD = 1.87$, 95% BC_a CI = (–3.67, –0.75), $z = -1.90$, $p = 0.111$, Hedge's $g = -0.95$].

We extracted data from the two individual-magnitude conditions in Experiment 2a and from the two intermixed conditions in Experiment 2b to explore the possible role of mislocation errors. A linear mixed effects model with the fixed effects of flanker phase perturbation and response set composition (with or without flanker presence) and random effect of observers was fitted to the log (base 10) threshold elevations (marginal $R^2 = 0.36$; conditional $R^2 = 0.64$). The omnibus test found a statistically significant main effect of flanker phase perturbation, $F(1, 31) = 24.10$, $p < 0.001$, but not of response set composition, $F(1, 9) = 3.24$, $p = 0.106$, on threshold elevation. The interaction between phase perturbation and response set composition was also statistically significant, $F(1, 31) = 5.11$, $p = 0.031$.

Discussion

Blocking the trials of intact-phase and phase-perturbed flanked conditions attenuated the difference in crowding strength between the two. However, the phase-perturbation-by-trial-blocking interaction was not statistically significant despite the nonsignificant planned comparison between the flanked intact-phase and flanked perturbed-phase conditions when the flanked trials were blocked. The large effect size of phase perturbation under the blocked condition (Hedge's $g = -0.95$) suggested a lack of statistical power due to small sample size as the reason for not achieving statistical significance. Therefore, we interpreted the results of Experiment 2b as a replicated effect of phase perturbation on crowding strength in Chinese-character recognition. Perturbed-phase flankers led to weakened crowding even when possible mislocation errors in intact-phase trials had been controlled for. Target–flanker similarity in contour integrity can modulate crowding strength in Chinese-character recognition.

Although phase perturbation weakened crowding strength in our experiments, crowding of Chinese characters remained significant (i.e., threshold elevation > 1) even with phase-perturbed flankers. Ho and Cheung (2011) found that crowding in an orientation discrimination task remained significant even when the flankers were rendered invisible through continuous flash suppression. Nonetheless, invisible flankers resulted in weaker crowding. Crowding by invisible flankers suggested a mechanism toward the early stage of visual processing, but the visibility modulation of crowding strength pointed to another late-stage component. Our current observation of weakened but significant crowding by phase-perturbed flankers was consistent with the idea that crowding happens at multiple stages of visual processing.

We also observed that trial blocking resulted in stronger crowding in general. Uncertainties in flanker type might have drawn spatial attention toward the whole trigram when the flanked trials were randomly intermixed. Such overall increase in deployed spatial attention might have attenuated crowding strength. Although attention was not the focus of our study, our findings suggested further investigation on the role of spatial attention in modulating crowding strength.

General discussion

Target–flanker similarity at the stroke-/contour-level modulates the strength of crowding in Chinese-character recognition, but similarity at the stroke-pattern level does not. Crowding remains strong as long as the general form of the flankers is maintained. In other words, high target–flanker similarity at the stroke-/contour-level facilitates faulty integration or pooling of features from target and flankers. Parkes et al. (2001) argued that pooling could happen as compulsory averaging of (low-level) visual features from target and flankers. However, some also suggested that pooling was a result of failed attentional selection (Whitney & Levi, 2011). One possible explanation for our observed similarity effects on crowding is that flankers similar to the target in general form can be harder to segment and distinguish from the target. According to S. He, Cavanagh, and Intriligator (1996), crowding is a result of limited spatial resolution in visual attention. Dissimilar flankers may result in pop-out of the target, a well-documented phenomenon in visual attention and visual search (Treisman, 1985; Wolfe, 1994). A target crowded by dissimilar flankers may overcome the limited attentional resolution through pop-out and therefore is relieved from strong crowding. An alternative explanation for the stroke-level similarity effect on crowding is that the features being pooled between target and flankers in Chinese-character crowding are the strokes. The quest for the responsible features in crowding can be a crucial step toward understanding the mechanism behind crowding. Previous studies on object crowding and face crowding have shed some light on this question.

Crowding of Chinese characters, objects, faces and English words

Wallace and Tjan (2011) found that object crowding was similar to letter crowding in terms of the radial–tangential anisotropy and the spatial extent of crowding. Such shared properties in crowding could be due to the features involved being the same in object crowding and letter crowding. Narrowband features could be a candidate. Chung et al. (2001) found that crowding of bandpassed letters was strongest when the flanking letters were filtered by a bandpass filter with a similar (but slightly lower) spatial frequency band. Features involved in crowding can be determined by both the spatial-frequency characteristics of the target, the demand of the task, and the properties of the visual system. Our current findings on Chinese-character crowding suggest that such narrowband features are closely related to the strokes and contours in the stimuli.

Although strokes and contours seem to play an important role in crowding of letters, objects, and Chinese characters, research on face crowding had shown that object-level representation could modulate crowding as well. Louie et al. (2007) found that upright face flankers resulted in stronger crowding than inverted face flankers when the target was an upright face. Holistic representation in upright faces might have modulated crowding strength. Louie et al.'s findings are consistent with crowding happening at multiple stages (Ho & Cheung, 2011). However, Sun and Balas (2015) could not replicate such flanker inversion effect on crowding of upright faces. The role of flanker inversion in face crowding is yet to be clarified.

Is holistic processing unique to face recognition? Martelli, Majaj, and Pelli (2005) compared the recognition of English letters within a word to that of facial features within a face. Recognition of both letters and facial features in foveal vision benefited from having a context. Such context superiority effect at fovea indicated similar object-level representations as words or faces was involved. Interestingly, context led to worse recognition performance in peripheral vision suggesting internal crowding, i.e., crowding from other letters in the word or other facial features in the face. Word recognition was similar to face recognition in both foveal superiority and peripheral inferiority effects. Despite the evidence of word-level representation, internal configuration of the flanker words did not affect the speed of reading a series of flanked words in a rapid serial visual presentation (RSVP) paradigm (Yu, Akau, & Chung, 2012). Yu et al. (2012) found that reading speed of the flanked RSVP stream was similar across different flanker conditions, including normal words, scrambled words, words with rotated letters, and words of different mirror image forms. Both our current findings on Chinese-character crowding and Yu et al.'s findings are consistent with the notion that strokes or contours are closely related to the features being misintegrated in crowding.

Crowding in reading

The important implications of crowding in reading have motivated decades of research since Bouma's (1970) seminal work on letter crowding. We make fixations with saccades in between as we read. Accurate letter recognition happens only in a small field around the fixation. Bouma called that the "functional visual field in reading" (p. 177). More specifically, Legge and colleagues defined the "visual span" in reading as the number of characters that an observer can recognize at each glance (Legge, Ahn, Klitz, & Luebker, 1997; Legge et al., 2007; Legge, Mansfield, & Chung, 2001). Visual-span size was shown to be tightly linked with reading speed (Legge et al., 2007). Pelli et al. (2007) further argued that the visual span was solely defined by crowding and called it the "uncrowded span." Although the visual process underlying English reading have been studied for decades, relatively little is known about Chinese reading. Research on Chinese-character crowding will be a crucial step in understanding the mechanisms behind Chinese reading.

Conclusions

Target–flanker similarity in orthographic legality does not affect Chinese-character crowding, but similarity in contour integrity does. Faulty integration of low-level visual features may be a key contributor in this. Being a class of objects with a large range of visual complexity familiar to more than one billion Chinese readers, the study of Chinese-character recognition and crowding can provide important insights into how the visual system identifies learned visual objects.

Keywords: crowding, Chinese-character recognition, faulty integration, strokes, phase perturbation, visual features

Acknowledgments

This research was supported in part by grants from the Research Grants Council, Hong Kong (General Research Fund—HKU 741213H and HKU 17676216) and in part by an internal grant from the University of Hong Kong (201209176178) to S.-H. C. Preliminary findings were presented as a poster at the Vision Sciences Society Annual Meeting 2015.

Commercial relationships: none.
Corresponding author: Sing-Hang Cheung.
Email: singhang@hku.hk.
Address: Department of Psychology, The University of Hong Kong, Hong Kong.

References

Anderson, E. J., Dakin, S. C., Schwarzkopf, S., Rees, G., & Greenwood, J. A. (2012). The neural correlates of crowding-induced changes in appearance. *Current Biology*, *22*(13), 1199–1206, doi:10.1016/j.cub.2012.04.063.

Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research*, *16*(1), 71–78, doi:10.1016/0042-6989(76)90078-X.

Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, *53*(6), 452–471, doi:10.1037/h0044049.

Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2007). Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cerebral Cortex*, *17*(7), 1604–1611, doi:10.1093/cercor/bhl071.

Bernard, J.-B., & Chung, S. T. L. (2011). The dependence of crowding on flanker complexity and target-flanker similarity. *Journal of Vision*, *11*(8):1, 1–16, doi:10.1167/11.8.1. [PubMed] [Article]

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177–178, doi:10.1038/226177a0.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436, doi:10.1163/156856897X00257.

Canty, A., & Ripley, B. (2016). *boot: Bootstrap R (S-Plus) Functions*. R package version 1.3-18. Retrieved from https://CRAN.R-project.org/package=boot

Chakravarthi, R., & Pelli, D. G. (2011). The same binding in contour integration and crowding. *Journal of Vision*, *11*(8):10, 1–12, doi:10.1167/11.8.10. [PubMed] [Article]

Chung, S. T. L., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research*, *41*(14), 1833–1850, doi:10.1016/S0042-6989(01)00071-2.

Chung, S. T. L., & Mansfield, J. S. (2009). Contrast polarity differences reduce crowding but do not benefit reading performance in peripheral vision. *Vision Research*, *49*(23), 2782–2789, doi:10.1016/j.visres.2009.08.013.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*(397), 171–185, doi:10.2307/2289144.

Freeman, J., Donner, T. H., & Heeger, D. J. (2011). Inter-area correlations in the ventral visual pathway reflect feature integration. *Journal of Vision*, *11*(4):15, 1–23, doi:10.1167/11.4.15. [PubMed] [Article]

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the central stream. *Nature Neuroscience*, *14*(9), 1195–1201, doi:10.1038/nn.2889.

Glen, J. C., & Dakin, S. C. (2013). Orientation-crowding within contours. *Journal of Vision*, *13*(8): 14, 1–11, doi:10.1167/13.8.14. [PubMed] [Article]

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences, USA*, *106*(31), 13130–13135, doi:10.1073/pnas.0901352106.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2012). Crowding follows the binding of relative position and orientation. *Journal of Vision*, *12*(3):18, 1–20, doi:10.1167/12.3.18. [PubMed] [Article]

He, C., & Tjan, B. S. (2004). What crowds a letter in the periphery? *Journal of Vision*, *4*(8): 508, doi:10.1167/4.8.508. [Abstract]

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature*, *383*(6598), 334–337, doi:10.1038/383334a0.

Ho, C., & Cheung, S.-H. (2011). Crowding by invisible flankers. *PLoS One*, *6*(12), e28814, doi:10.1371/journal.pone.0028814.

Hong Kong SAR Education Bureau. (2009). *Lexical items with English explanations for fundamental Chinese learning in Hong Kong schools*. Hong Kong: Education Bureau, The Government of the Hong Kong Special Administrative Region.

Hong Kong SAR Education Bureau. (2012). *List of graphemes of commonly used Chinese characters*. Hong Kong: Education Bureau, The Government of the Hong Kong Special Administrative Region.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363, doi:10.1002/bimj.200810425.

Hsiao, J. H., & Cottrell, G. W. (2009). Not all visual expertise is holistic, but it may be leftist. *Psychological Science*, *20*(4), 455–463, doi:10.1111/j.1467-9280.2009.02315.x.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*(2), 255–279, doi:10.1163/156856894X00350.

Lefcheck, J. S. (2015). piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579, doi:10.1111/2041.210X.12512.

Legge, G. E., Ahn, S. J., Klitz, T. S., & Luebker, A. (1997). Psychophysics of reading—XVI. The visual span in normal and low vision. *Vision Research*, *37*(14), 1999–2010, doi:10.1016/S0042-6989(97)00017-5.

Legge, G. E., Cheung, S.-H., Yu, D., Chung, S. T. L., Lee, H.-W., & Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of Vision*, *7*(2):9, 1–15, doi:10.1167/7.2.9. [PubMed] [Article]

Legge, G. E., Mansfield, J. S., & Chung, S. T. L. (2001). Psychophysics of reading: XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, *41*(6), 725–743, doi:10.1016/S0042-6989(00)00295-9.

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*(5), 635–654, doi:10.1016/j.visres.2007.12.009.

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

Louie, E. G., 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, 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, doi:10.1167/5.1.6. [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, doi:10.1167/7.2.5. [PubMed] [Article]

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Research*, *32*(4), 771–777, doi:10.1016/0042-6989(92)90192-L.

Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, *69*(5), 529–541, doi:10.1109/PROC.1981.12022.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*(7), 739–744, doi:10.1038/89532.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442, doi:10.1163/156856897X00366.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*(28), 4646–4674, doi:10.1016/j.visres.2006.04.023.

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, doi:10.1167/4.12.12. [PubMed] [Article]

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, *7*(2):20, 1–36, doi:10.1167/7.2.20. [PubMed] [Article]

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2016). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-128. Retrieved from https://CRAN.R-project.org/package=nlme

Põder, E. (2007). Effect of colour pop-out on the recognition of letters in crowding conditions. *Psychological Research*, *71*(6), 641–645, doi:10.1007/s00426-006-0053-7.

Poon, W. Y., & Kang, B. W. (2003). *A study of the Chinese characters recommended for the subject of Chinese language in primary schools*. Hong Kong: Language Centre, Faculty of Arts, Hong Kong Baptist University.

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Shin, K., Wallace, J. M., & Tjan, B. S. (2010). Objects crowded by noise flankers. *Journal of Vision*, *10*(7):1336, doi:10.1167/10.7.1336. [Abstract]

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*(3), 471–477, doi:10.1016/0002-9394(62)94878-X.

Sun, H.-M., & Balas, B. (2015). Face features and face configurations both contribute to visual crowding. *Attention, Perception, & Psychophysics*, *77*(2), 508–519, doi:10.3758/s13414-014-0786-0.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*(7), 1349–1357, doi:10.1016/0042-6989(92)90227-A.

Torchiano, M. (2016). *effsize: Efficient Effect Size Computation*. R package version 0.7.0. Retrieved from https://CRAN.R-project.org/package=effsize.

Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, *31*(2), 156–177, doi:10.1016/S0734-189X(85)80004-9.

van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology*, *6*(1), e1000646. doi:10.1371/journal.pcbi.1000646.

Wallace, J. M., & Tjan, B. S. (2011). Object crowding. *Journal of Vision*, *11*(6):19, 1–17, doi:10.1167/11.6.19. [PubMed] [Article]

Wang, Z., & Simoncelli, E. (2004). Local phase coherence and the perception of blur. In S. Thrun, L. Saul, & B. Schölkopf (Eds.), *Advances in neural information processing systems* (pp. 786–792). Cambridge, MA: MIT Press.

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120, doi:10.3758/BF03202828.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168, doi:10.1016/j.tics.2011.02.005.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238, doi:10.3758/BF03200774.

Yu, D., Akau, M. M. U., & Chung, S. T. L. (2012). The mechanism of word crowding. *Vision Research*, *52*(1), 61–69, doi:10.1016/j.visres.2011.10.015.

Zhang, J.-Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2007). Legibility variations of Chinese characters and implications for visual acuity measurement in Chinese reading population. *Investigative Ophthalmology & Visual Science*, *48*(5), 2383–2390. [PubMed] [Article]

Zhang, J.-Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2009). Legibility of Chinese characters in peripheral vision and the top-down influences on crowding. *Vision Research*, *49*(1), 44–53, doi:10.1016/j.visres.2008.09.021.