The dependence of crowding on flanker complexity and target-flanker similarity

Jean-Baptiste Bernard and
School of Optometry, Vision Science Graduate Program, University of California, Berkeley, CA 94720

Susana T.L. Chung
School of Optometry, Vision Science Graduate Program, Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720

Abstract

We examined the effects of the spatial complexity of flankers and target-flanker similarity on the performance of identifying crowded letters. On each trial, observers identified the middle character of random strings of three characters ("trigrams") briefly presented at 10° below fixation. We tested the 26 lowercase letters of the Times-Roman and Courier fonts, a set of 79 characters (letters and non-letters) of the Times-Roman font, and the uppercase letters of two highly complex ornamental fonts, Edwardian and Aristocrat. Spatial complexity of characters was quantified by the length of the morphological skeleton of each character, and target-flanker similarity was defined based on a psychometric similarity matrix. Our results showed that (1) letter identification error rate increases with flanker complexity up to a certain value, beyond which error rate becomes independent of flanker complexity; (2) the increase of error rate is slower for high-complexity target letters; (3) error rate increases with target-flanker similarity; and (4) mislocation error rate increases with target-flanker similarity. These findings, combined with the current understanding of the faulty feature integration account of crowding, provide some constraints of how the feature integration process could cause perceptual errors.

Keywords

crowding; spatial vision; letter recognition; object recognition; reading

INTRODUCTION

The ability to recognize objects in peripheral vision is severely limited when the object of regard is surrounded in close proximity by other objects. This is the crowding effect (Bouma, 1970; see Levi, 2008 and Pelli & Tillman, 2008 for review). Crowding occurs for simple targets such as short line segments or Gabor patches (Andriessen & Bouma, 1976; He, Cavanagh & Intriligator, 1996; Parkes, Lund, Angelucci, Solomon & Morgan, 2001; Pöder, 2008; Pöder & Wagemans, 2007; Wilkinson, Wilson & Ellemberg, 1997) and also for more complex stimuli such as alphanumeric characters (Bouma, 1970; Chung, Levi & Legge, 2001; Pelli, Palomares & Majaj, 2004; Strasburger, Harvey & Rentschler, 1991), words (Chung, 2004) and faces (Louie, Bressler & Whitney, 2007; Martelli, Majaj & Pelli, 2005). It has been suggested that crowding occurs when flankers are presented within an *integration field*[1] around the target (Toet & Levi, 1992; Pelli et al, 2004). Despite the ample

Corresponding author: Jean-Baptiste Bernard, School of Optometry, University of California, Berkeley, CA 94720-2020, jb.bernard@berkeley.edu.

evidence demonstrating the degrading effect of flankers within this integration field, the mechanism that underlies how flankers degrade the target signal during crowding remains unclear. Our current knowledge of crowding indicates that crowding affects the identification of the details of the target, but not the detection of the target (Chung, 2010; Levi, Hariharan & Klein, 2002; Pelli et al, 2004), and that crowding is distinct from pattern masking (Chung et al., 2001; Levi et al, 2002; Pelli et al, 2004) and surround suppression (Petrov, Popple, & McKee, 2007). Some properties of a crowded target, such as spatial frequency and orientation, although cannot be identified, are still available to the visual system as they remain capable of inducing an adaptation effect (He et al, 1996), contribute to the computation of the averaged signal pooled across the crowded target and its flankers (Parkes et al, 2001), and the computation of more complex statistics (Balas, Nakano & Rosenholtz, 2009).

Many theories have been suggested to account for crowding, ranging from the optics of the eye, receptive field size to attention (see Levi, 2008, for a review). Although currently there is no consensus as to which one is *the* mechanism underlying crowding, a popular account of the crowding effect is that features (the most basic elements detected by the visual system) from the target and flankers are combined inappropriately at a stage subsequent to the detection of the features and prior to the identification of the target, thus giving rise to a non-veridical percept of the target (Levi et al, 2002; Nandy & Tjan, 2007; Pelli et al, 2004; Pelli & Tillman, 2008; Van Den Berg, Roerdink & Cornelissen, 2010). Despite the popularity of this *faulty feature integration* account of crowding, it remains unclear as to how or why features of the target and flankers are combined inappropriately or incorrectly. Pelli et al. (2004) postulated that crowding is due to excessive integration, a consequence of the absence of integration fields that are of an appropriate size in the periphery, and thus features from the target and flankers are integrated over an inappropriately larger integration field. This idea is similar to the suggestion that the second-stage filter or the template used for identifying a target beyond the detection stage is not well matched to the crowded target (Levi, Klein & Carney, 2000; Levi, McGraw & Klein, 2000). Using a novel classification image paradigm and by analyzing the noise pattern along with the flanker signal for each trial, Nandy and Tjan (2007) found that crowding reduces the amount of valid features (defined as the proportion of target features used by human observers that are also utilized by an ideal observer) while increases the amount of invalid features used by the visual system. This finding implies that the excessive integration of features over an inappropriately large integration field may not be a compromise due to the lack of smaller integration fields in the periphery (Pelli et al, 2004), but instead, is a necessity in order for the visual system to get enough valid features for integration. A prediction that can be made based on the supposition of Pelli et al (2004) and Nandy and Tjan (2007) is that crowding should increase with the number of flankers or flanker features within the integration field.

How crowding increases with the number of flankers has been investigated in several recent studies. In general, the magnitude of crowding or the number of error for identifying crowded objects increases with the number of flankers around a target (Levi & Carney, 2009; Põder & Wagemans, 2007), except when the additional flankers are placed along the tangential direction with respect to the fovea, which has less crowding than the radial direction (Pelli et al, 2004; Wilkinson et al, 1997) or when the additional flankers can be grouped to form a contextual configuration (Levi & Carney, 2009; Livne & Sagi, 2007; 2010). These findings are consistent with the prediction of excessive integration within the integration field as the cause of crowding based on two accounts. First, an increase in the

---

[1]*Also known by different terminologies such as the isolation field (Pelli et al, 2004), integration zone (Levi et al, 2002), perceptive hypercolumn (Levi, Aitsebaomo & Klein, 1985), combining field (Pelli & Tillman, 2008) etc. In this paper, we will adopt the term integration field.*

number of flankers directly leads to an increase in the number of flanker features that could be integrated with the target features, thus allowing for more possible combinations between target and flanker features and the different global percept of the resulting combinations. The visual system would then have to choose one global percept among many alternatives. Based simply on probability, the error rate of identifying the target would increase. Second, an increase in the number of flankers could simply increase the number of invalid features present in the integration field. However, to our knowledge, there is no published data on the direct effect of the number of flanker *features* on crowding. This could be attributed to the fact that it is often difficult to define what a feature is, especially for more complicated objects like alphanumeric symbols (Grainger, Rey & Dufau, 2008). To circumvent this difficulty, Pelli, Burns, Farell & Moore-Page (2006) proposed that the number of features in an alphanumeric symbol is proportional to its spatial *complexity*, which has since been quantified in several ways, including the perimetric complexity and stroke frequency. Perimetric complexity is defined as the square of the inside-and-outside perimeter of a symbol, divided by the "ink" area (Arnoult & Attneave, 1956; Pelli et al, 2006). Stroke frequency refers to the average number of lines crossed by a slice through a symbol, divided by the symbol width (Majaj, Pelli, Kurshan & Palomares, 2002). In its simplest form, slices through a symbol were only made along the horizontal direction (Majaj et al, 2002). A variation of this stroke frequency measurement is to take into account that complexity may not be evenly distributed across the symbol, especially for more complex symbols such as Chinese characters. To compute the stroke frequency of Chinese characters, Zhang, Zhang, Xue, Liu & Yu (2007; 2009) positioned the slices horizontally on the upper and lower halves of each character, vertically on the left and right halves, and along oblique orientations (±45° to horizontal). The maximum number of crossings made by these slices is taken as the stroke frequency. The *first* goal of this study was to investigate the dependence of crowding on the spatial complexity of letters, which we assumed as a surrogate of the number of features as proposed by Pelli et al (2006). We focused our investigation on the characters of the Roman alphabet because of our interest in understanding the limiting factors on English reading, with the ultimate goal of developing strategies or reading materials that could enhance reading speed.

It is well known that crowding increases with target-flanker similarity of shape (Kooi, Toet, Triparthy & Levi, 1994; Nazir, 1992), orientation (Andriessen & Bouma, 1976; Levi et al, 2002), spatial frequency (Chung et al, 2001), contrast polarity (Chung & Mansfield,2009; Kooi et al, 1994), color (Kooi et al, 1994; Pöder, 2007) or visual complexity (Zhang et al, 2009). Very often, the reduction in crowding when the target and flankers are dissimilar is attributed to top-down factors, such as pop-out or attention (Kooi et al,1994; Levi, 2008). Here we hypothesized that the effect of target-flanker similarity on crowding could also be explained based on the faulty feature integration hypothesis. In particular, we hypothesized that when a target and a flanker share common or similar features, the features from the flanker that are not shared with the target might be integrated more readily with those of the target, thus causing confusion. Alternatively, the features from the flanker that are similar to those of the target could compete with the target features for the integration process as suggested by Nandy and Tjan (2007). Based on these considerations, the *second* goal of this study was to examine the dependence of crowding on the target-flanker similarity. To do so, we first devised a quantitative psychometric measurement of letter similarity based on a modified confusion matrix constructed from human observations, and then examined how crowding depends on the target-flanker similarity. Here, we assumed that the similarity score between two letters is directly related to the number of common or similar features shared by the two letters.

To preview our results, we found that crowding increases with flanker complexity up to a certain value, after which crowding does not depend on flanker complexity. We also found that crowding increases with target-flanker similarity.

METHODS

Experiment 1

We analyzed data collected from two previous studies in which observers were trained to identify crowded letters in peripheral vision (Chung, 2007; Truong, Arunkumar, Yu & Chung, 2009). Eight young observers with normal vision (aged 18–31; best-corrected acuity in each eye 20/20 or better) participated in each study. Although the purposes of the two studies were different and there were some methodological differences between the two studies, data used for our analyses were obtained using fundamentally identical procedures. In both studies, on each trial, the observers' task was to identify a target letter flanked closely by two other letters, one on the right and the other on the left, at 10° eccentricity in the inferior visual field. All three letters were randomly drawn from the 26 lowercase letters of the Roman alphabet, and were presented together as an ensemble ("trigram"). The center-to-center separation between the target and its flanking letter was fixed at 0.8× the *x*-height, a spacing closer than the standard spacing used in ordinary text (approximately 1.16× for Courier font: Chung, 2002). This separation yielded substantial crowding effect (see Results). Each observer was tested with 1000 trials each day, for six consecutive days (6000 trials total). In the study of Chung (2007), the font used was Times-Roman, and each trigram was presented for 150 ms at a letter size equivalent to 1.4× of each observer's critical print size for reading. Although Times-Roman is a proportional-width font, in that study, letters in each trigram were rendered at a fixed letter separation. In the study of Truong et al (2009), Courier was used as the font and each trigram was presented for 100 ms at a letter size corresponding to 1.4× the letter size that yielded 52% correct identification for single letters. Averaged across the eight observers in each study, the letter size used was 1.80° (range: 1.34° to 2.30°) and 1.31° (range: 0.91° to 1.68°) in Chung (2007) and Truong et al (2009), respectively. As we shall see later, despite the differences in the stimuli between the two studies, the main findings were qualitatively the same. The inclusion of data obtained using two different fonts allowed us to test the generality of our hypothesis.

Calculation of spatial complexity

We used the length of the morphological skeleton of a letter to represent its spatial complexity (see Figure 1 for the skeletons of the letter *g*). The morphological skeleton of a binary shape represents the centers of the maximal disks located inside the shape. The advantages of using such a measurement include its independence of the boldness of the letter stroke, thus facilitating comparison across fonts, and its simplicity and intuitiveness in relation to the definition of spatial complexity of a letter. A comparison of the length of the skeleton measurement with the perimetric complexity for the entire set of 26 letters yields a correlation coefficient of 0.95 for the Times-Roman font and 0.90 for the Courier font, showing a strong correlation between these measurements.

To create the skeleton of a letter, we first created templates for each letter of the Times-Roman and Courier fonts for a letter size (specified as the *x*-height) of 43 pixels. This size is large enough to ensure that details of the letters would be represented faithfully and not limited by the spatial resolution. We used software custom-written in MATLAB (Mathworks, MA; version 7.1) to construct the morphological skeleton of each letter template and then calculated the length of the skeleton in pixels. Table 1 presents the spatial complexity scores, expressed as the length of the skeleton of each letter normalized by the *x*-height, for Times-Roman (mean: 136.12 ± 35.87 [SD]; range: 85 to 222) and Courier (mean:

152.15 ± 24.68 [SD]; range 118 to 194). Like perimetric complexity (Pelli et al, 2006) and stroke frequency (Majaj et al, 2002), our measure of spatial complexity in terms of the skeleton length is additive. Thus, we can express the flanker complexity of a trigram as the sum of the complexity scores of the two flankers.

Measurement of similarity between two letters

To quantify the similarity between two letters, we defined a metric based on a conventional confusion matrix[2] with the main difference being that trials for which observers reported one of the flankers as the target (mislocation errors) were excluded. The exclusion of mislocation error trials is necessary because mislocation errors could arise as a result of letter position uncertainty (Chung & Legge, 2009), which could contaminate the target-flanker similarity score. To construct the similarity matrix, we took all 48,000 trials (6000 trials from each observer) for each font, and excluded the trials in which mislocation errors occurred. Mislocation errors amounted to 15% and 14% for Times-Roman and Courier fonts, respectively. Then for each font, we constructed a 26×26 matrix with the stimulus letters (middle letters of trigrams) represented in rows, from *a* to *z*, and the responses of observers represented in columns. Each cell in this matrix represents the probability P_{ij} that a letter I_i was identified as I_j, which we adopted as the similarity score between a pair of letters. As for values in a conventional confusion matrix, the similarity score for a given pair of letters is asymmetric. In relation to the second goal of this paper, we sought to determine if a higher similarity between a target and its flanking letters, implying more common features shared by the target and flankers, increases the strength of crowding. For this analysis, we discarded trials in which at least one of the flankers was identical to the target. Then, we expressed the target-flanker similarity of a trigram as the *sum* of the similarity scores between the target letter and each of its flankers. For example, the similarity score $C(F_1/T/F_2)$ for a trigram composed of the three letters F_1, T and F_2 is $P_{TF_1} + P_{TF_2}$. We chose to use the sum of the similarity scores because we hypothesized that the effect of target-flanker similarity is directly related to the number of features shared by the target and the flankers, and thus is additive in nature.

Additional Experiments

As we shall see later, the general results for the Times-Roman and Courier fonts are similar, raising a question of whether the results are truly generalizable across different fonts, or whether the results represent the mere effect that the characters for the two fonts are too similar so that the spatial complexity score of a letter is essentially indicative of the identity of the letter, regardless of the font. After all, the spatial complexity of the 26 letters spans a range of 85 to 222 for Times-Roman and 118 to 194 for Courier font, which are relatively similar. To address the question of whether or not our findings are truly generalizable to other fonts or characters, we measured the performance for identifying crowded characters that spanned a larger range of spatial complexity scores. We extended the range of spatial complexity scores using two approaches. First, we examined the intra-font spatial complexity effect by including non-letter symbols, lower- and uppercase letters of the same font (Experiment 2). Second, we used two ornamental fonts with intricate details (Experiment 3).

Experiment 2—In this experiment, we tested five observers with normal vision (aged: 19–37) using a similar experimental paradigm as in Chung (2007) but instead of using only the 26 lowercase Times-Roman letters as stimuli, we used a total of 79 symbols/optotypes

---

[2]Here, our similarity scores between two letters were derived from a confusion matrix obtained with crowded letters. In a recent study, Freeman et al (2010) derived letter similarity scores based on a confusion matrix obtained with single letters, and showed qualitatively similar results as ours with respect to the effect of letter similarity on the rate of mislocation errors.

(referred to as *characters* in this experiment). These 79 characters included 27 non-letter symbols (see Table 1 for the list of symbols used), 26 lowercase and 26 uppercase of the Times-Roman font. The spatial complexity score for this entire character set ranges from 17 (for the symbol '.')[3] to 300 (for the letter '*W*'). On each trial, the target and the flankers were randomly drawn from the 79 characters. Each trigram was presented for 100 ms at 10° in the inferior visual field. The character size (*x*-height) ranged between 0.6 and 1.2° (average = 0.94°) for the five observers. To avoid overlapping between characters, especially for the uppercase letters, we used a character spacing of 1.1× the *x*-height, slightly larger than that used in Chung (2007). Each observer was tested with 4000 trials. For comparison, each observer also identified single, unflanked characters for 100 trials.

Experiment 3—Five observers with normal vision (aged: 19–32) were tested using a similar experimental paradigm as described in Experiments 1 and 2. Three of these observers also participated in Experiment 2. Characters were the 26 uppercase letters of the Edwardian and Aristocrat fonts. Spatial complexity was determined using a character size such that the height of the uppercase letter *X* of the Edwardian and Aristocrat fonts (determined separately) matched that of the Times and Courier fonts (43 pixels). For the Edwardian font, the spatial complexity scores ranged from 192 ('*𝒥*') to 492 ('*ℳ*'), while the spatial complexity scores for the Aristocrat font ranged from 331 ('*𝒪*') to 646 ('*𝒢*'). Table 1 shows spatial complexity values for all the characters used in this experiment. In any given block of trials, the target and flankers were either all drawn from the Edwardian or the Aristocrat font.

As in the Experiments 1 and 2, each trigram was presented for 100 ms at 10° in the inferior visual field. Across the five observers, the character size (the height of the uppercase *X*) ranged between 0.6 and 1.2° (average = 0.88°) for both fonts. The letter spacing was 1.75× the X-height to avoid overlapping between two adjacent letters, while at the same time small enough to produce crowding. Each observer completed 4000 trials (2000 trials for each font). Because these fonts were unfamiliar to our observers, prior to data collection, observers practiced the task of recognizing these characters presented singly for 200 trials at the fovea, followed by 400 trials at 10° eccentricity in the inferior visual field. At the end of the practice session, observers' performance for recognizing these single characters was above 80% correct (the last 100 trials) for both fonts. The set of uppercase letters was permanently displayed on a second computer screen, so that observers could refer to these letters when needed.

RESULTS

Experiment 1

Effect of crowding on letter identification—Averaged across the 48,000 trials, the proportion-correct of identifying crowded letters was 0.56 (SD: 0.07; range: 0.48 – 0.64 across observers) for Times-Roman font in the study of Chung (2007). This performance was substantially lower than that for identifying single, unflanked letters of the same size (proportion-correct: 0.95 or higher, for the same eight observers), demonstrating a substantial crowding effect. Similarly, the proportion-correct of identifying crowded letters for Courier font averaged 0.50 (SD: 0.12; range: 0.40 – 0.73 across observers) in the study of Truong et al (2009). Although Truong et al did not measure the performance for identifying single, unflanked letters, the averaged performance of their observers for

---

[3]In theory, the complexity score for the period symbol should be 1. However, using our Matlab function to create the template, the period symbol was not a perfect round disc because of pixelation. Consequently, the morphological skeleton of the period symbol was not simply a dot (or a pixel) which has a theoretical complexity score of 1.

identifying flanked letters at a larger letter separation (2× *x*-height) ranged between 0.85 and 0.92, demonstrating a substantial crowding effect. Note that these data were collected from perceptual learning studies, and that the recognition rate varied on average from 0.48 (the first 1000 trials of each observer) to 0.62 (the last 1000 trials) for the Times-Roman font, and from 0.45 (the first 1000 trials) to 0.56 (the last 1000 trials) for the Courier font.

We attribute the errors in identifying crowded letters to three main sources: the similarity between the target and the response letters, mislocation errors and random errors (e.g. finger errors and response bias). Our analyses showed that approximately one-third (31 ± 6% [SD] for Times-Roman and 32 ± 7% [SD] for Courier) of all the errors made were due to mislocation. As for random errors, there is no easy way to calculate the rate of their occurrence, but based on observers' reports, we assumed that the rate for random errors was negligible.

Effect of the complexity of flankers on crowding—The proportion-incorrect (error rate) of identifying crowded target letters in trigrams is plotted as a function of the complexity score of flankers (sum of the individual complexity scores for the two flankers) for Times-Roman and Courier fonts in Figure 2. In each panel, different symbols represent the data for each of the eight observers tested with the respective font. To facilitate the visualization of the effect of flanker complexity on the error rate of identifying crowded letters, we binned the flanker complexity scores into 11 complexity intervals, starting from 161 to 400, in intervals of 20. Each interval contains at least 83 trials from each observer. The adoption of 11 complexity intervals represented a good compromise between having enough intervals to demonstrate a relationship between error rate and flanker complexity and to contain sufficient number of trials in each interval. For both fonts and for all observers, the error rate increased with flanker complexity. The relationship between the error rate and the flanker complexity score can be approximated by a linear regression function, as shown in Figure 2c, where the averaged error rates across the eight observers for each font are plotted. The change of error rate for the same range of complexity score was faster for Times Roman than for Courier, as indicated by the difference in the slopes (0.14 vs. 0.09%-error per unit of complexity score) of the regression functions.

Figure 2 shows that the error rate of identifying crowded target letters in trigrams increased with the complexity of flankers. Did the error rate also depend on the *target* complexity? Here, we divided the target complexity into three groups for each font — low, medium and high, so that each group contained approximately equal number of letters. As shown in Figure 3, the increase of error rate with flanker complexity depended on target complexity. Error rate increased faster with flanker complexity when the target complexity was low and slower when the target complexity was high. Another way of describing this result is that target letters of low complexity (such as *c, i, l*) were more susceptible to the complexity of the flankers than their higher-complexity counterparts (such as *g, m, w*).

Effect of the target-flanker similarity on crowding—Matrices summarizing observers' responses to crowded target letters for Times-Roman and Courier fonts, excluding trials for which mislocation errors occurred, are presented in Table 2, where stimulus and response letters are shown in rows and columns, respectively. A given cell in the matrices (i, j) represents the averaged probability that the stimulus letter l_i was identified as l_j. For example, for Times-Roman font, when the letter *k* was presented as the middle letter in trigrams, there was a probability of 1.67% that it was mistaken as the letter *x* when letter *x* was not one of the flankers. On the other hand, when letter *x* was presented as the middle letters in trigrams, the probability that it was identified as letter *k* was 6.07%. We refer to the matrices as the *psychometric similarity matrices* and the probability in each cell as a similarity score. Clearly, the similarity scores are directional. Between the matrices for

the two different fonts, the similarity scores for the same pair (and direction) of target-flanker are generally very similar, despite some differences in the physical aspects of the letters.

The second prediction of this paper was that the similarity of a target letter and its flankers could explain, in large part, the errors made in identifying crowded target letters. To test our prediction, we calculated the target-flanker similarity score for each trial (the sum of the similarity score between the target and each of its flankers). For example, for the trigram *edc* rendered in Times-Roman, the target-flanker similarity score for the letter pair *ed* when letter *d* was the stimulus letter is 1.09%, and the similarity score for the letter pair *dc* when the letter *d* was the stimulus letter is 2.40%, yielding an overall similarity score of 3.49% for the trigram. Across all trials, the target-flanker similarity scores averaged 3.13% (range: 0% – 53.46%) for the Times-Roman font, and 3.60% (range: 0% – 60.53%) for the Courier font. We binned the similarity scores into 9 intervals. Each interval contains at least 27 trials for Times Roman and 42 trials for Courier from each observer. Figure 4 shows the error rate for identifying crowded letters as a function of the target-flanker similarity score for each of the eight observers and for the two fonts. In this figure, the target-flanker similarity scores on the x-axis are plotted in a logarithmic scale because the similarity scores are not uniformly distributed across the range; instead, most of the scores clustered around low values. Averaged across observers, the increase in error rate with the target-flanker similarity score followed a linear function on the semi-log plot (Figure 4c). These findings imply that the similarity between the target and its flanking letters also plays an important role in causing identification errors of crowded letters. Interestingly, the rate of mislocation errors, in proportion to the total number of trials, also increased linearly with the target-flanker similarity score on a semi-log plot (Figure 5). The significance of the similar relationships between the total error rate and the rate of mislocation errors with letter similarity will be discussed in Discussion.

Experiment 2

To determine if the results shown in Figure 2 on how crowding increases with flanker complexity applies to characters of a larger range of spatial complexity within the same font, we tested five observers using a set of 79 characters of the Times-Roman font with spatial complexity scores ranging from 17 to 300. Averaged across observers, the performance accuracy for identifying *single* characters was 0.9, and fell to 0.58 (range: 0.44 – 0.70) for identifying crowded characters. The unfilled blue circles in Figure 6 show that the error rate of identifying crowded characters increased with flanker complexity, akin to what we observed in Figure 2. We binned the data into complexity intervals with the same interval width as in Figure 2, with a total of 22 complexity intervals (a minimum of 20 trials per observer). For comparison, results obtained for using only the 26 lowercase letters (from Figure 2) are included in this figure. Clearly, regardless of the set size (26 *vs.* 79) and the range of spatial complexity scores, error rate increased with flanker complexity, suggesting that the effect could be generalized to characters that span a larger range of spatial complexity.

However, the interacting effect of target and flanker complexity on crowding that we observed earlier in Figure 3 does not generalize to this set of data obtained with a larger range of complexity (see Figure 7a). As in Figure 3, we divided all the trials into three groups (low, medium and high) according to the target complexity score. Although targets of high complexity scores were still the least susceptible to the effect of flanker complexity (the change of error rate with flanker complexity was the slowest), there was very little difference on the effect of flanker complexity for the low and medium target complexity groups. The implication of this finding will be discussed in Discussion.

Experiment 3

To further extend the range of spatial complexity scores to higher values, we tested two ornamental fonts with intricate details — Edwardian and Aristocrat. These data are presented in Figure 6. For comparison, results from Figure 2 for lowercase letters of the Times-Roman (filled blue circles) and Courier (filled red squares) are replotted in this figure. Filled brown diamonds and green triangles represent data for the Edwardian and Aristocrat font, respectively. We binned the data into complexity intervals starting from 1 to 1200 in intervals of 100, with a minimum of 20 trials from each observer. Clearly, the relationships between error rate of character identification vs. flanker complexity for the different fonts all collapsed into one single function, where the error rate increased with flanker complexity until it reached approximately 60%, beyond which error rate became independent of flanker complexity. This result indicates that although flanker complexity is an important factor limiting crowding, it cannot completely account for crowding. When we divided the trials into three groups according to target complexity (low, medium and high), there was very little difference on the effect of flanker complexity for the different target complexity groups.

DISCUSSION

The purpose of this paper was to understand how spatial complexity and target-flanker similarity contribute to letter crowding. Our results show that (1) the error rate of identifying flanked letters increases with the spatial complexity of the flankers up to a certain value, beyond which error rate becomes independent of flanker complexity; (2) target letters of higher complexity are less affected by flankers than target letters of lower complexity; (3) the error rate of identifying flanked letters increases with target-flanker similarity; and (4) the mislocation error rate increases with target-flanker similarity. The combined effects of spatial complexity of flankers and target-flanker similarity can account for a large range of letter recognition performance.

Effect of spatial complexity on letter crowding

Figures 2 and 6 show that the error rate for identifying crowded characters increases with the spatial complexity of the flankers, up to a certain value, beyond which the error rate does not depend on further increase in spatial complexity. We interpret the spatial complexity independent region as a saturation effect.

According to our supposition (following that of Pelli et al, 2006) that spatial complexity is a surrogate for the number of features within a character, an increase in the spatial complexity of the flankers implies that there are more features contributed by the flankers within the integration field. An increase in the number of flanker features increases the number of possible ways that target and flanker features could be combined, thus increasing the chance for a perceptual error. This explanation applies even if we consider that some of the flanker features are invalid because the visual system apparently uses more invalid features during crowding (Nandy & Tjan, 2007), and that an increase in the total number of flanker features would proportionally increase the number of invalid features.

In our analysis shown in Figures 2 and 6, we divided the trials into different groups according to the sum of the complexity measurements of the right and left flankers. Because many combinations of flankers could yield a similar sum of flanker complexity, the relatively small error bars in Figures 2 and 6 provide evidence that it is the total *number* of flanker features within the integration field, instead of a particular combination of flanker features, that is an important determinant of crowding.

Besides flanker complexity, target complexity also plays a role in determining the crowding effect. Figure 3 shows that target letters of low complexity are more susceptible to the flanker effect, whereas target letters of high complexity are more immune to the flanker effect. This finding is an interesting contrast to the result for single-letter identification. Pelli et al (2006) showed that human efficiency for letter identification is higher for low complexity letters, and lower for high complexity letters. Here, we show that in the presence of flanking letters, letter identification performance is better for high complexity target letters than for low. Our finding does not contradict that of Pelli et al. Instead, it provides some insight into the feature integration process. There is evidence that during the feature integration process, features derived from different parts of the visual field are not weighted the same; instead, they are weighted according to their spatial position. For example, it is well known that letter flankers affect target identification more when they are closer to the target — a classical characteristic of the crowding effect. The distance-dependent flanker effect can be interpreted as a higher probability for a feature to be included for integration if it is closer to the target. Recently, van den Berg et al (2010) proposed a model that is based on the "distributed population coding" of neuronal processing (Zemel, Dayan & Pouget, 1998) to explain feature integration during crowding. According to their model, the weighting assigned to each feature in the integration field depends on the cortical distance between the center of the integration field and the neuron that encodes the given feature. Features that are close to the center of the integration field receive higher weights than features that are farther away, and have a higher probability of being integrated. In their model, they assumed the weighting follows a 2D Gaussian profile.

In relation to our study, if we assume that the integration field is centered on the target letter, and that there is no mislocation or displacement of letter features due to spatial uncertainty, then features of the target letter would be given more weight than features of the flanking letters during the integration process. A target letter of low complexity will therefore contribute fewer highly weighted features to the center of the integration field, making it easier for the feature integration process to be disrupted by flanker features present in the integration field, especially if there are a larger number of flanker features and that they are close to the target. This logic would also apply in the presence of positional uncertainty of the letter features (target and flankers).

Nandy and Tjan (2007) argued that feature integration is essentially a competitive process. According to their argument, very often there are different ways the visual system can integrate visual information to form a global percept, and that the visual system chooses one over the other(s). Although we still do not have a good handle on how the visual system makes its decision, presumably the presence of a large number of flanker features could easily compete with the target features, and bias the visual system as to which features to integrate. An important factor, therefore, is the proportion of the target and flanker features within the integration field. A target letter of low complexity implies a lower proportion of target features relative to flanker features, and thus the flanker features could better "compete" or bias the visual system in integrating these features, causing perceptual errors.

Interestingly, the systematic interactions between target and flanker complexity on identification error rate that we observed in Figure 3 seem to break down when the stimulus set comprises a large range of complexity (see Figure 7). This result could be explained by at least two possibilities. In an earlier paragraph, we suggested that whether a feature (from the target or flanker) will be included in the feature integration process depends solely on its weighting, which in turn, depends on the cortical distance between the neuron that encodes the feature and the center of the integration field, and can be subjected to spatial uncertainty. If there is a "threshold" for the minimum number of features from the target that need to be included, or the minimum weighting that will be assigned to the target features, then even

for a very simple symbol such as a dot, its susceptibility to the flanker effects would be diminished compared with an integration process based straightly on a weighting scheme. Alternatively, the weaker interaction effect between target and flanker complexity on the identification error rate could be explained as a similarity effect. As we shall explain in greater details in the next section, when two characters are similar, the common features shared by the two characters may serve as "anchors" for the dissimilar features to be attached and ultimately integrated to form a final percept of the stimulus character. For two characters that do not share any common features, such as when they have very different complexity scores (e.g. a period and the letter *W*), there is no common feature to serve as the anchor and thus the chance that the flanker features are integrated with some of the target features to produce an erroneous percept of the target is diminished. In fact, this explanation also accounts for the findings of Zhang et al (2009) who found that crowding was reduced when target and flankers were derived from different sets of Chinese characters with very different spatial complexities.

When the spatial complexity of the flankers is very high, the error rate of identifying crowded characters shows a saturation effect, which occurred at an error rate of approximately 60%. This result indicates that flanker complexity is not the only factor that limits identification performance. For instance, target-flanker similarity could also limit identification performance (see Figure 4). The saturation effect implies that when there is a large number of irrelevant or invalid features within the integration field, the addition of extra irrelevant features is not going to change the proportion of irrelevant features involved in the integration process. In other words, there may exist a maximum number of flanker features that could be integrated.

Effect of target-flanker similarity on letter crowding

Previous studies investigating the effect of target-flanker similarity on crowding often specified similarity by dichotomizing some physical characteristics of the target and flankers, such as color (Kooi et al, 1994; Pöder, 2007) and shape (Kooi et al, 1994; Nazir, 1992). In this study, we defined target-flanker similarity based on the psychometric similarity matrix (a modified confusion matrix with mislocation trials excluded). Some earlier studies have taken the approach of using a confusion matrix to examine target-flanker similarity. For example, Krumhansl and Thomas (1977), Chastain (1982; 1985) each derived a small set (usually the set size = 2) of similar and dissimilar letters based on the scores in the confusion matrix of Townsend (1971), and examined how the error rate of identifying a target letter depended on whether the flanker (or *masker* according to their terminology) was from the same (similar) or a different (dissimilar) set of letters as that of the target. Interestingly, the results from these three studies were vastly different. Krumhansl and Thomas (1977) found a small effect of degraded letter identification performance when the pair of letters was similar (more confusable) compared with a pair of dissimilar letters, with the effect occurring only at a short duration (65 ms) but not at a longer duration (90 ms). Similarly, Chastain (1982) found that the error rate of identifying a target letter was higher when the flanker was similar to the target and lower when the flanker was dissimilar to the target. However, in his 1985 paper, he reported the opposite effect — the error rate of identifying a target letter was higher when the target and the flanker were from different letter sets. Further, Townsend and Snodgrass (1977) reported a null effect of letter similarity on letter identification accuracy. They did not rely on a confusion matrix to select a pair of similar and a pair of dissimilar letters. Instead, they defined similarity based on similar morphological structures shared by letters. For instance, letters *O* and *D* share a curved structure, letters *I* and *L* share a vertical structure. Clearly, from these papers, the effect of target-flanker similarity on letter identification performance was inconclusive. None of these studies used more than four letters in a given experiment (two similar and two dissimilar

letters) and the four letters used were different in all the studies. If the effect was specific to the set of four letters in each study, then it could explain why the effect of target-flanker similarity differed among studies. In our study, we used the entire set of 26 lowercase letters, therefore, the generalizability of the effect of target-letter similarity would not be a concern.

Our finding of an increase in letter identification error rate with target-flanker similarity can be the consequence of several possibilities. As mentioned earlier, flanker features may compete with target features to be included in the feature integration process. If so, it is likely that the more similar a flanker feature is to one of the target features, the higher the probability that it will be included if the resulting percept is likely to be closer to a real letter. Alternatively, our result can also be explained by feature mislocation, a direct consequence of spatial uncertainty. The idea is similar to the model proposed by Chung and Legge (2009) in which they assumed that the encoded position of each letter is independent and Gaussian distributed, and that the distribution of the spread governs the precision of localizing a letter. In their study, the spatial uncertainty was with respect to the whole letter; here, we argue that spatial uncertainty could occur at the feature level. For a pair of target and flanker that share some common features, spatial uncertainty in localizing the *common* features would not affect how features are perceived and subsequently combined to form the percept of a letter. These common features may serve as the "anchors" for other features to be combined. However, spatial uncertainty in localizing features that are not shared between the target and flankers could cause the incorrect features from the flankers to be combined with the "anchors", thus causing perceptual errors. An important assumption for this supposition is that after the features are combined, there is a decision stage or "template-matching" at which the resultant combination of features is compared against the templates of all the possible letters. According to this supposition, it is almost impossible for a pair of dissimilar letters (example: *o* and *x*) to produce perceptual error because even when features are mislocalized due to spatial uncertainty, the resultant combinations of features do not resemble any of the 26 real letters and thus these combinations of features would be rejected by the visual system at the decision stage. For letters that are similar, perceptual errors increase with target-flanker similarity (more common features) because there is an increased likelihood that the resultant combination of features would yield a valid percept. It is also likely that when a target letter and its flankers share many similar features (a high target-flanker similarity score), fewer flanker features would need to be mislocalized to be combined with the anchors to form a percept that is close to a real letter and thus result in a perceptual error. Even if a target and its flankers share no common feature, it is still possible that some of the flanker features become mislocalized and get combined with the target features to form a percept of a letter that is unlike the target or the flankers. However, in our analysis, this type of identification error is not easily distinguishable from just random errors. Therefore, for now, we will simply focus on the scenarios where the target and flanking letters share some similar features. Based on our results (see also Freeman et al, 2010; Pöder & Wagemans, 2007), we speculate that the mislocation errors during crowding are mostly due to feature mislocations and subsequent combinations between target and flanker features within the integration field.

Figure 4 shows the relationship between letter identification error rate and target-flanker similarity. Although the trials in which mislocation errors occurred were excluded in the derivation of the target-flanker similarity score, the data were obtained from the same observers and our argument might appear circular. To ensure that the relationship between identification error rate and target-flanker similarity score was not because these measurements were derived from the same data-set, we applied the same target-flanker similarity matrix to another set of 48,000 trials of data that were not included in the original analyses. These 48,000 trials were obtained from a different group of eight observers in the

study of Truong et al (2009). As shown by the dashed line in Figure 4c, the target-flanker similarity score derived from our original analysis could predict the overall letter identification error rate of a different group of eight observers, implying a robust relationship between letter identification error rate and our measure of target-flanker similarity.

Our finding that crowding increases with flanker complexity seems to be at odds with a recent study examining how the legibility of Chinese characters is affected by flankers of similar or very different complexity (Zhang et al, 2009). In their study, Zhang et al used two groups of Chinese characters that have very different complexity scores, and found that the size threshold for recognizing a Chinese character was higher in the presence of flanking characters — the conventional crowding effect. However, the increase in size threshold was larger when the target and flanking characters belonged to the same complexity category and smaller when the target and flanking characters belonged to different complexity categories, independent of whether the target character was of a low or high complexity. In other words, the results of Zhang et al (2009) imply that crowding depends on the similarity or difference in complexity between the target and flankers, instead of showing a monotonic dependence on flanker complexity (at least up to a certain value), as we found in this study (Figures 2 and 6). The difference between our result and Zhang et al's could be explained as follows. Recall that a large proportion of letter identification errors is due to feature mislocation and that in order for feature mislocation to occur, the target and its flanker need to share some similar features. When a target and its flankers were drawn from very different complexity categories, as in Zhang et al (2009), feature mislocation could not occur and thus there was less crowding. In essence, Zhang et al's findings could be explained as a similarity effect, and not a complexity effect. In fact, if we accept that Zhang et al's findings is due to similarity, then their results are reminiscent of that of Chastain's (1982) who defined similarity based on confusability scores whereas Zhang et al (2009) defined similarity based on spatial complexity.

Interaction between complexity and target-flanker similarity on crowding

We show that both letter complexity and target-flanker similarity can modulate the crowding effect.[4] In Figure 8 we plot the letter identification error rate as a function of flanker complexity and target-flanker similarity for the Courier and Times-Roman fonts. We divided the ranges of flanker complexity and target-flanker similarity scores each into eight intervals. When both factors are considered, identification error rate could be as low as 16% for Courier and 10% for Times-Roman, and could reach as high as 75% for Courier and 89% for Times-Roman. These large ranges illustrate the importance of considering both letter complexity and target-flanker similarity in understanding how letter recognition is affected by the presence of nearby letters.

Considering that we used a fixed center-to-center separation to render letters in the trigrams, could our complexity result be explained by a target-flanker separation effect? For instance, a flanker *w* is closer of the target letter than the letter *l*. Figure 3 shows that low-complexity target letters (e.g. *i, l*) constitute the group most susceptible to the flanker effect, compared with other groups, and the target-flanker separation is usually larger for these low-complexity target letters. Therefore, we do not think that the complexity result is due to the fixed letter separation used in the experiment.

Summary: "Rules" for feature integration within an integration field

Based on our findings, we propose several "rules" as to how features from objects in close proximity are integrated within an integration field. These "rules" provide constraints for

any quantitative models that attempt to explain how feature integration could lead to perceptual errors.

1. Within the integration field, features from the target and flankers have equal chance to be integrated if everything else is the same (but see below). In other words, the information as to whether a particular feature comes from the target or flanker is lost. This idea is similar to the Principle of Univariance for photon absorption, that once a photon is photoisomerized, the information about the wavelength of the photon is lost.

2. Features of the target and flankers are weighted within an integration field (assumed to be centered on the target) according to the cortical distance between the neuron that encodes a feature and the center of the integration field. Features that are closer to the center of the integration field are weighted higher than those further away from the center of the integration field. Spatial uncertainty of features could give rise to feature mislocation and render a flanker feature that is further from the center of the target to appear as if it is closer to the center of the integration field, and thus receives a higher weighting.

3. There exists a threshold for the minimum number of target features that need to be included, or the minimum weighting that the target features carry, in the feature integration process (evidence provided by Figure 7).

4. There is a maximum number of flanker features that could be included in the feature integration process (evidence provided by Figure 6).

5. Within an integration field, the probability of a feature being integrated is directly related to the frequency of occurrence of that feature, and is independent of whether this feature belongs to the target and/or the flanker.

6. Common features shared by a target and a flanker will serve as "anchors" for other features to be combined. A pair of letters that do not share any common features will not have any anchoring features and thus even in the presence of feature mislocation, perceptual errors will not occur.

Acknowledgments

REFERENCES

Andriessen JJ, Bouma H. Eccentric vision: adverse interactions between line segments. Vision Research. 1976; 16:71–78. [PubMed: 1258390]

Arnoult MD, Attneave F. The quantitative study of shape and pattern perception. Psychological bulletin. 1956; 53:452–471. [PubMed: 13370691]

Balas B, Nakano L, Rosenholtz R. A summary-statistic representation in peripheral vision explains visual crowding. Journal of Vision. 2009; 9(12):1–18.

Bouma H. Interaction effects in parafoveal letter recognition. Nature. 1970; 226:177–178. [PubMed: 5437004]

Chastain G. Confusability and interference between members of parafoveal letter pairs. Perception & Psychophysics. 1982; 32:576–580. [PubMed: 7167357]

Chastain G. Masking effects with targets and nontargets of varying similarity. Acta Psychologica. 1985; 60:1–10. [PubMed: 4050515]

Chung STL. The effect of letter spacing on reading speed in central and peripheral vision. Investigative Ophthalmology & Visual Science. 2002; 43:1270–1276. [PubMed: 11923275]

Chung STL. Reading speed benefits from increased vertical word spacing in normal peripheral vision. Optometry and Vision Science. 2004; 81:525–535. [PubMed: 15252352]

Chung STL. Learning to identify crowded letters: does it improve reading speed? Vision Research. 2007; 47:3150–3159. [PubMed: 17928026]

Chung STL. Detection and identification of crowded mirror-image letters in normal peripheral vision. Vision research. 2010; 50:337–345. [PubMed: 19961868]

Chung STL, Legge GE. Precision of position signals for letters. Vision Research. 2009; 49:1948–1960. [PubMed: 19460396]

Chung STL, Levi DM, Legge GE. Spatial-frequency and contrast properties of crowding. Vision Research. 2001; 41:1833–1850. [PubMed: 11369047]

Chung STL, Mansfield JS. Contrast polarity differences reduce crowding but do not benefit reading performance in peripheral vision. Vision Research. 2009; 49:2782–2789. [PubMed: 19695281]

Freeman, J.; Pelli, DG.; Chakravarthi, R. Paper submitted. 2010. Compulsory pooling of crowded objects.

Grainger J, Rey A, Dufau S. Letter perception: from pixels to pandemonium. Trends in cognitive sciences. 2008; 12:381–387. [PubMed: 18760658]

He S, Cavanagh P, Intriligator J. Attentional resolution and the locus of visual awareness. Nature. 1996; 383:334–337. [PubMed: 8848045]

Kooi FL, Toet A, Tripathy SP, Levi DM. The effect of similarity and duration on spatial interaction in peripheral vision. Spatial Vision. 1994; 8:255–279. [PubMed: 7993878]

Krumhansl CL, Thomas EAC. Effect of level of confusability on reporting letters from briefly presented visual arrays. Perception & Psychophysics. 1977; 21:269–279.

Levi DM. Crowding — an essential bottleneck for object recognition: a mini-review. Vision Research. 2008; 48:635–654. [PubMed: 18226828]

Levi DM, Aitsebaomo AP, Klein SA. Vernier acuity, crowding and cortical magnification. Vision Research. 1985; 25:963–977. [PubMed: 4049746]

Levi DM, Carney T. Crowding in peripheral vision: Why bigger is better. Current Biology. 2009; 19:1988–1993. [PubMed: 19853450]

Levi DM, Hariharan S, Klein SA. Suppressive and facilitatory spatial interactions in peripheral vision: peripheral crowding is neither size invariant nor simple contrast masking. Journal of Vision. 2002; 2(2):167–177. [PubMed: 12678590]

Levi DM, Klein SA, Carney T. Unmasking the mechanisms for Vernier acuity: evidence for a template model for Vernier acuity. Vision Research. 2000; 40:951–972. [PubMed: 10720666]

Levi DM, McGraw PV, Klein S. Vernier and contrast discrimination in central and peripheral vision. Vision Research. 2000; 40:973–988. [PubMed: 10720667]

Livne T, Sagi D. Configuration influence on crowding. Journal of Vision. 2007; 7(2):4, 1–12. [PubMed: 18217819]

Livne T, Sagi D. How do flankers' relations affect crowding? Journal of Vision. 2010; 10(3):1–14. [PubMed: 20377278]

Louie EG, Bressler DW, Whitney D. Holistic crowding: Selective interference between configural representations of faces in crowded scenes. Journal of Vision. 2007; 7(2):24, 1–11. [PubMed: 18217839]

Majaj NJ, Pelli DG, Kurshan P, Palomares M. The role of spatial frequency channels in letter identification. Vision Research. 2002; 42:1165–1184. [PubMed: 11997055]

Martelli M, Majaj NJ, Pelli DG. Are faces processed like words? A diagnostic test for recognition by parts. Journal of Vision. 2005; 5(1):58–70. [PubMed: 15831067]

Nandy AS, Tjan BS. The nature of letter crowding as revealed by first- and second-order classification images. Journal of Vision. 2007; 7(2):5, 1–26. [PubMed: 18217820]

Nazir TA. Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. Vision Research. 1992; 32:771–777. [PubMed: 1413560]

Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. Compulsory averaging of crowded orientation signals in human vision. Nature Neuroscience. 2001; 4:739–744.

Pelli DG, Burns CW, Farell B, Moore-Page DC. Feature detection and letter identification. Vision Research. 2006; 46:4646–4674. [PubMed: 16808957]

Pelli DG, Palomares M, Majaj NJ. Crowding is unlike ordinary masking: distinguishing feature integration from detection. Journal of Vision. 2004; 4(12):1136–1169. [PubMed: 15669917]

Pelli DG, Tillman KA. The uncrowded window of object recognition. Nature Neuroscience. 2008; 11:1129–1135.

Petrov Y, Popple AV, McKee SP. Crowding and surround suppression: Not to be confused. Journal of Vision. 2007; 7(2):12, 1–9. [PubMed: 18217827]

Pöder E. Effect of colour pop-out on the recognition of letters in crowding conditions. Psychological Research. 2007; 71:641–645. [PubMed: 16718511]

Pöder E. Crowding with detection and coarse discrimination of simple visual features. Journal of Vision. 2008; 8(4):24, 1–6. [PubMed: 18484863]

Pöder E, Wagemans J. Crowding with conjunctions of simple features. Journal of Vision. 2007; 7(2):3, 1–12.

Strasburger H, Harvey LO Jr. Rentschler I. Contrast thresholds for identification of numeric characters in direct and eccentric view. Perception and Psychophysics. 1991; 49:495–508. [PubMed: 1857623]

Toet A, Levi DM. The two-dimensional shape of spatial interaction zones in the parafovea. Vision Research. 1992; 32:1349–1357. [PubMed: 1455707]

Townsend G. Theoretical analysis of an alphabetic confusion matrix. Perception & Psychophysics. 1971; 9:40–50.

Townsend G, Snodgrass JG. Masking effects of similar and dissimilar letters: A test of the interactive channels model. Perceptual and Motor Skills. 1977; 44:1212–1228.

Truong S, Arunkumar A, Yu D, Chung STL. A comparison of the effectiveness of perceptual learning on a daily vs. weekly basis. Optometry & Vision Science. 2009; 86 E-abstract 90842.

Van Den Berg R, Roerdink JBTM, Cornelissen F. A neurophysiologically plausible population code model for feature integration explains visual crowding. PLoS Computational Biology. 2010; 6:1–10.

Wilkinson F, Wilson HR, Ellemberg D. Lateral interactions in peripherally viewed texture arrays. Journal of the Optical Society of America. 1997; 14:2057–2068. [PubMed: 9291601]

Zemel R, Dayan P, Pouget A. Probabilistic interpretation of population codes. Neural Computation. 1998; 10:403–430. [PubMed: 9472488]

Zhang JY, Zhang T, Xue F, Liu L, Yu C. Legibility variations of Chinese characters and implications for visual acuity measurement in Chinese reading population. Investigative Ophthalmology and Visual Science. 2007; 48:2383–2390. [PubMed: 17460306]

Zhang JY, Zhang T, Xue F, Liu L, Yu C. Legibility of Chinese characters in peripheral vision and the top-down influences on crowding. Vision Research. 2009; 49:44–53. [PubMed: 18929592]
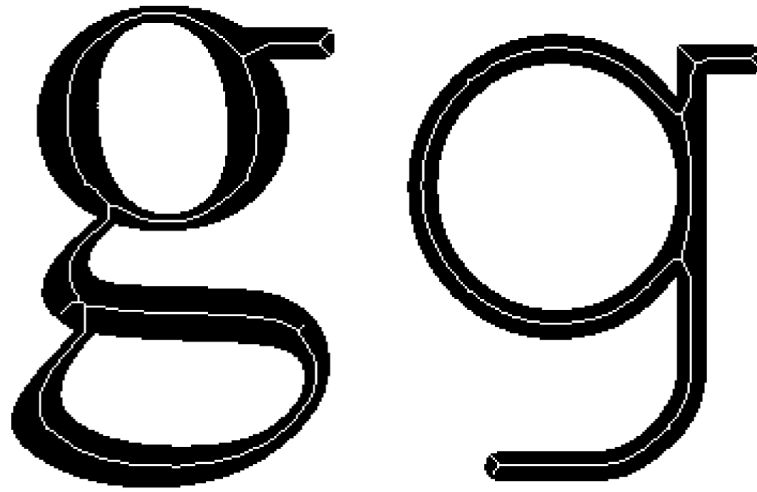
Figure 1.
The templates for the letter *g* and their skeletons as rendered in Times-Roman (left) and Courier (right) font. Spatial complexity of a letter is defined as the length (number of pixels) of the morphological skeleton.
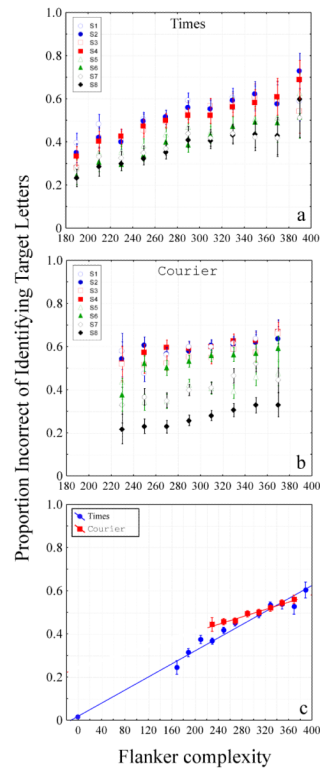
Figure 2.
Proportion-incorrect for identifying the middle (target) letters of trigrams is plotted as a function of flanker complexity for Times-Roman (a) and Courier (b) fonts. In each panel, different colored symbols represent data from individual observers. Data averaged across the eight observers for each font are shown in panel c. For Times-Roman, the datum plotted at a flanker complexity of 0 represents the unflanked condition. Note the change in scale on the x-axis for panel c. Error bars represent the 95% confidence interval.
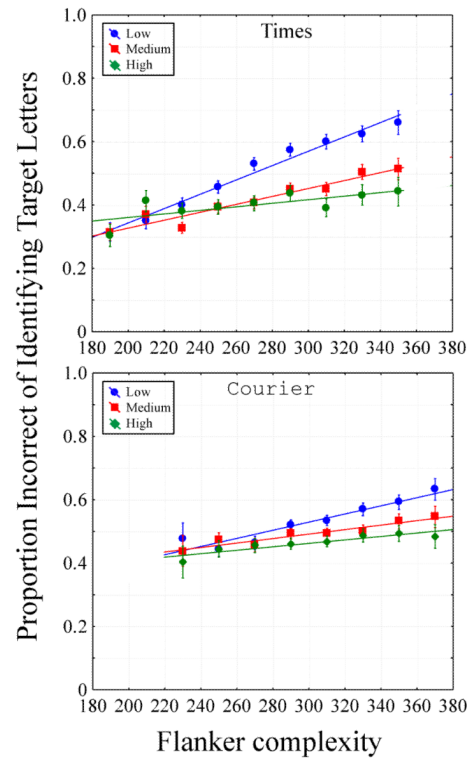
Figure 3.
Proportion-incorrect for identifying the middle (target) letters of trigrams is plotted as a function of flanker complexity for the three categories of target complexity, for Times-Roman (top) and Courier (bottom) font. The slopes for the regression lines are 0.26, 0.14 and 0.11 %-error per unit of complexity score for Times-Roman and 0.13, 0.07 and 0.05 %-error per unit of complexity score for Courier. Error bars represent the 95% confidence interval.

Figure 4.
Proportion-incorrect for identifying the middle (target) letters of trigrams is plotted as a function of the logarithm of the target-flanker similarity score for Times-Roman (a) and Courier (b) fonts. Details of the figure are as in Figure 2. Unfilled red squares ("Courier 2") represent the data also collected from the study of Truong et al (2009) but from a different group of eight observers (see text for details). Error bars represent the 95% confidence interval.
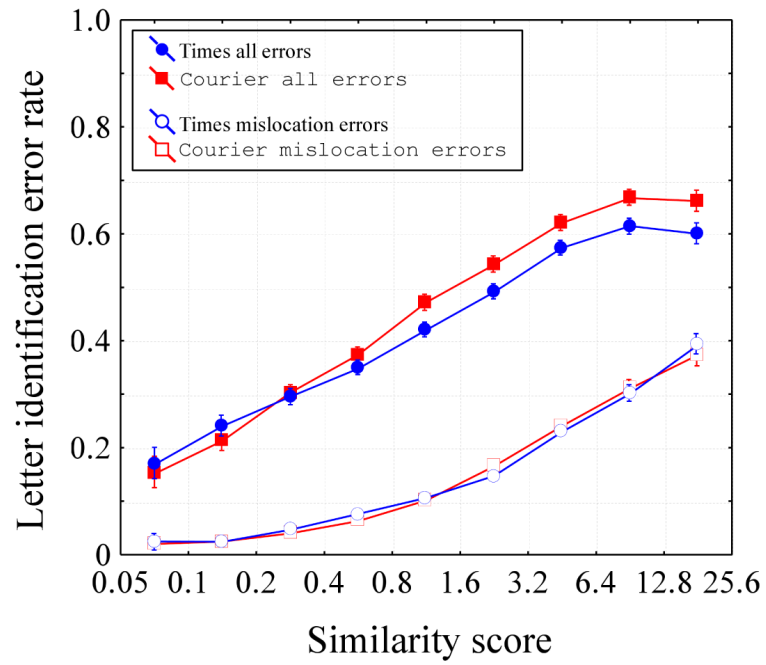
Figure 5.
Letter identification and mislocation error rate are plotted as a function of the logarithm of the target-flanker similarity for Times-Roman (a) and Courier (b) fonts.
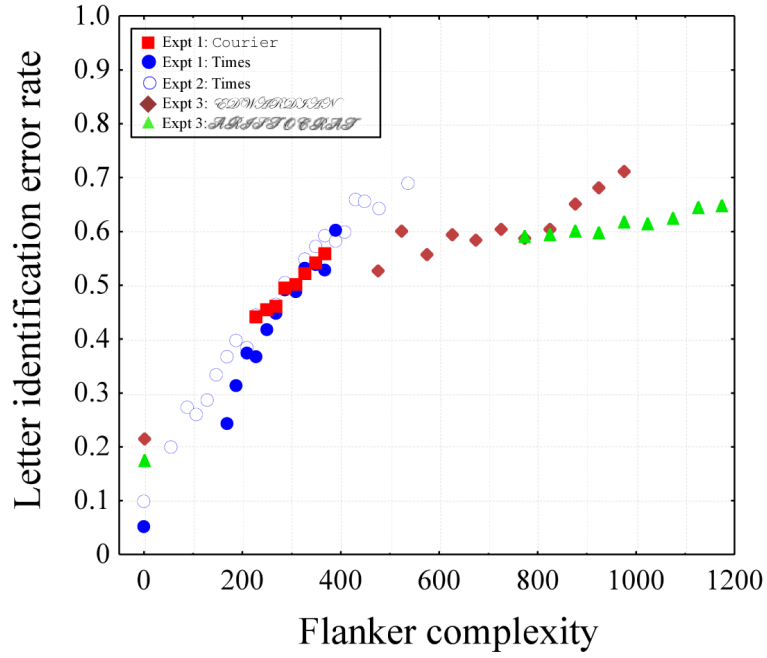
Figure 6.
Proportion-incorrect for identifying the middle (target) symbols of trigrams is plotted as a function of flanker complexity for all three experiments. Data points are color-coded and represent the different set of characters used (see legend for details). Data shown represent the averaged values across the group of observers tested for each font. Data points plotted at a flanker complexity of 0 represent the performance for the unflanked conditions.
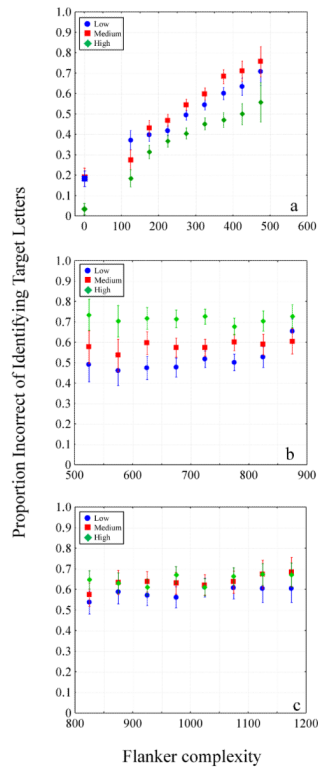
Figure 7.
Proportion-incorrect for identifying the middle (target) symbols of trigrams is plotted as a function of flanker complexity for the three categories of target complexity for Experiment 2 (a: Times-Roman) and Experiment 3 (b: Edwardian; c: Aristocrat).
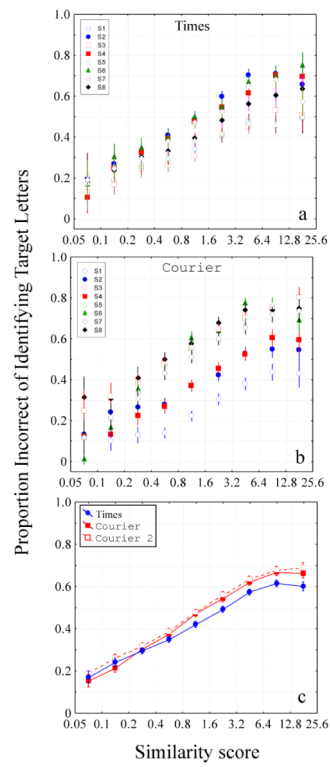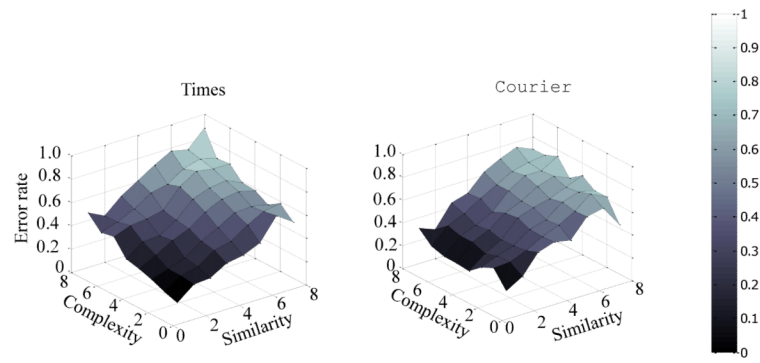
Figure 8.
Proportion-incorrect for identifying the middle (target) letters of trigrams is plotted as a function of flanker complexity and the logarithm of the target-flanker similarity for Times-Roman (left) and Courier (right) fonts. Data are binned in 64 different groups, according to their similarity (0.05 to 12.8%, in logarithmic intervals) and complexity scores (170 to 410, in intervals of 30 for the Times font and 220 to 380 in intervals of 20 for the Courier font). The shaded key on the right represents the different error rate.

Table 1

Courier[1]		Times[1,2]		Times[2]				Edwardian[3]		Aristocrat[3]	
a	156	a	140	A	185	!	66	A	398	A	513
b	177	b	149	B	225	#	232	B	460	B	545
c	119	c	95	C	151	$	221	C	339	C	553
d	176	d	161	D	199	%	226	D	402	D	532
e	154	e	121	E	230	^	70	E	355	E	478
f	159	f	137	F	180	&	235	F	383	F	549
g	180	g	192	G	196	*	118	G	458	G	646
h	173	h	165	H	253	(	91	H	402	H	604
i	120	i	85	I	107	)	91	I	191	I	387
j	131	j	105	J	103	-	32	J	216	J	407
k	163	k	173	K	225	_	54	K	349	K	539
l	120	l	90	L	148	=	108	L	375	L	462
m	194	m	222	M	300	+	108	M	492	M	635
n	150	n	143	N	222	[	119	N	355	N	499
o	118	o	108	O	165	{	102	O	270	O	331
p	191	p	165	P	169	]	119	P	367	P	530
q	188	q	157	Q	192	}	102	Q	310	Q	373
r	126	r	91	R	215	;	53	R	409	R	577
s	137	s	109	S	161	:	34	S	332	S	418
t	127	t	93	T	155	<<	96	T	315	T	539
u	141	u	131	U	188	,	35	U	331	U	390
v	130	v	109	V	159	<	103	V	292	V	376
w	178	w	195	W	275	.	17	W	415	W	479
x	155	x	141	X	212	/	74	X	428	X	428
y	161	y	133	Y	163			92	Y	397	Y	493
z	132	z	129	Z	176	?	100	Z	363	Z	467
						'	38				

Table 2

Stimulus	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
a		0.79	1.48	2.99	5.08	1.18	0.53	0.91	0.12	0.00	2.54	0.06	0.72	1.53	0.77	0.44	0.42	1.59	15.06	0.41	0.12	0.28	0.83	0.44	0.20	5.04
b	0.63		0.66	9.66	0.54	1.93	0.26	7.48	0.33	0.80	2.37	0.86	0.07	0.62	3.24	1.88	0.13	0.32	2.59	0.71	0.58	0.58	0.06	0.05	0.07	0.26
c	2.92	3.68		7.47	18.90	0.42	0.23	0.20	0.20	0.19	1.30	0.22	0.22	0.55	9.04	0.74	2.07	1.43	9.29	0.07	0.58	2.11	0.91	1.34	0.17	1.77
d	0.70	7.46	2.40		1.09	3.56	0.13	4.27	0.38	0.26	3.30	1.99	0.14	1.00	1.81	0.21	1.32	0.70	2.41	1.14	1.01	0.08	0.12	0.30	0.00	0.44
e	5.78	1.38	1.56	2.13		0.82	0.91	0.12	0.21	0.07	0.96	0.13	0.07	0.44	0.35	0.60	0.25	0.90	13.18	0.07	0.06	1.50	1.19	0.34	0.14	1.20
f	0.40	0.13	0.30	1.79	0.18		1.00	1.19	0.66	0.48	0.74	0.26	0.17	0.13	0.05	0.00	0.06	1.99	0.94	1.50	0.29	0.12	0.07	0.00	0.05	0.25
g	0.17	0.12	0.00	0.24	0.29	0.67		0.60	0.06	0.23	0.00	0.00	0.00	0.17	0.11	0.29	0.12	0.06	2.39	0.05	0.00	0.12	0.00	0.05	0.17	0.06
h	0.60	8.96	0.06	9.30	0.00	6.18	0.21		1.10	0.54	6.85	3.39	0.87	3.55	0.32	0.43	0.28	0.82	1.43	1.31	0.54	0.10	0.30	0.12	0.00	0.90
i	0.45	1.12	0.29	4.64	0.15	9.47	0.20	3.68		6.16	1.49	6.82	0.23	1.00	0.13	0.39	0.15	2.39	1.11	2.15	0.56	0.44	0.06	0.06	0.08	0.80
j	0.07	0.12	0.12	0.69	0.06	0.78	0.00	0.83	3.46		1.28	0.55	0.06	0.06	0.07	0.62	0.70	0.12	0.33	0.07	0.00	0.13	0.00	0.13	0.92	0.00
k	0.83	4.79	0.48	7.53	0.36	3.29	0.14	8.98	0.28	1.03		1.87	0.16	0.34	0.06	0.51	0.13	0.53	4.21	0.68	0.07	0.27	0.00	1.67	0.07	1.11
l	0.26	3.67	0.21	12.00	0.07	4.67	0.06	15.00	2.05	1.06	6.40		0.07	0.69	0.09	0.00	0.37	0.99	0.76	1.97	0.26	0.14	0.13	0.06	0.07	0.13
m	5.07	0.86	0.15	2.29	0.48	2.89	0.24	4.74	1.10	0.15	0.37	1.00		26.73	0.79	0.90	0.57	3.55	0.98	0.80	7.20	0.89	1.30	0.16	0.07	1.87
n	3.09	1.51	0.44	3.54	0.78	1.49	0.16	9.34	1.38	0.16	1.11	0.87	14.20		1.24	1.41	0.74	9.06	2.73	0.97	7.47	1.10	1.80	0.79	0.11	2.81
o	1.48	10.53	6.41	7.52	3.38	0.36	0.29	0.29	0.54	0.15	0.28	0.06	0.22	0.84		3.49	1.81	1.22	4.35	0.43	0.76	1.64	1.11	0.19	0.24	0.75
p	0.27	1.26	0.13	0.82	0.46	0.36	0.46	0.26	0.21	1.52	0.45	0.00	0.20	0.43	1.56		15.40	1.44	0.88	0.20	0.47	0.51	0.19	0.15	0.77	0.31
q	0.60	0.98	0.91	2.43	1.13	0.96	0.56	0.33	0.32	0.39	0.27	0.26	0.12	0.47	1.79	13.00		0.48	1.69	0.26	0.63	0.07	0.47	0.00	0.25	0.43
r	1.28	0.07	1.57	2.14	1.77	4.36	0.09	1.61	3.56	0.33	1.49	0.97	3.58	4.51	0.81	0.45	0.77		3.51	2.74	3.55	1.68	1.35	1.01	0.51	3.13
s	6.24	0.51	0.45	0.70	3.12	0.20	0.50	0.12	0.06	0.00	1.32	0.25	0.06	0.59	0.31	0.13	0.25	0.99		0.00	0.07	0.50	0.45	0.82	0.06	0.64
t	1.69	2.01	0.81	7.27	2.00	11.60	0.18	2.78	3.69	0.62	4.54	3.04	0.78	1.72	0.44	0.50	0.65	10.13	2.55		4.06	1.88	0.45	1.07	0.06	1.22
u	1.86	3.80	0.66	6.80	0.87	1.63	0.08	3.31	1.98	0.44	1.11	1.18	3.86	8.97	1.66	0.97	1.47	5.85	3.41	3.99		2.63	1.99	0.55	0.19	1.76
v	1.28	0.07	0.82	1.12	1.25	1.20	0.07	0.59	0.55	0.08	1.23	0.27	0.57	0.84	0.24	0.28	0.36	2.60	1.77	0.83	1.57		23.28	2.81	5.45	2.89
w	1.44	0.60	0.13	0.74	1.06	0.20	0.06	0.19	0.19	0.07	1.08	0.28	0.74	0.32	0.06	0.19	0.67	1.08	1.40	0.34	0.41	11.88		1.20	2.74	1.25
x	3.92	0.45	1.04	1.01	1.95	0.34	0.09	0.57	0.07	0.13	6.07	0.16	0.51	0.81	0.46	0.23	0.43	1.85	6.89	0.21	0.49	3.48	9.21		1.02	8.90
y	0.20	0.28	0.12	0.18	0.18	0.06	0.23	0.12	0.00	0.74	0.41	0.24	0.06	0.11	0.00	0.47	0.53	0.19	0.99	0.19	0.72	2.54	0.75	0.36		0.97
z	1.48	0.58	0.39	0.39	0.68	0.69	0.05	0.23	0.12	0.00	0.53	0.12	0.23	0.48	0.07	0.14	0.00	0.77	1.24	0.36	0.12	0.83	0.51	2.89	0.11	

Response

Table 3

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
a		1.16	0.13	1.86	9.35	1.89	0.86	1.14	0.51	0.00	1.33	1.67	0.87	0.55	0.78	1.07	0.98	0.26	21.40	0.34	0.20	0.37	1.40	1.12	0.49	3.08
b	0.87		0.35	19.90	0.23	2.19	1.72	8.24	1.67	1.40	4.16	5.74	0.62	0.43	3.11	2.84	0.08	0.20	1.37	1.26	1.01	0.28	0.15	0.22	0.08	0.24
c	3.59	6.47		11.20	2.45	2.69	2.93	1.00	0.40	0.74	1.09	0.36	0.28	1.04	13.70	3.45	3.09	3.01	5.94	0.62	1.32	1.34	0.71	1.10	0.72	1.85
d	1.22	10.80	1.20		0.14	4.92	2.04	3.78	1.25	0.72	2.35	4.49	0.14	0.70	2.14	1.56	1.29	0.68	1.49	1.63	1.67	0.14	0.07	0.44	0.07	0.14
e	14.00	2.09	0.73	1.81		1.46	0.86	0.59	0.00	0.31	0.71	0.26	0.12	0.15	0.61	1.68	0.45	0.51	30.30	0.12	0.07	0.59	0.33	0.79	0.14	2.72
f	0.25	0.24	0.17	1.06	0.25		0.18	1.20	0.19	0.00	0.55	2.80	0.00	0.12	0.05	0.00	0.00	0.58	0.25	1.78	0.00	0.13	0.05	0.19	0.00	0.19
g	0.35	0.87	0.00	0.62	0.00	0.32		0.43	0.00	4.56	0.13	0.21	0.21	0.18	0.42	4.40	4.92	0.20	0.71	0.14	0.47	0.06	0.00	0.36	3.21	0.15
h	0.45	7.32	0.21	6.23	0.07	6.44	0.43		2.84	1.39	7.34	17.20	1.87	1.75	0.56	0.73	0.35	0.16	0.66	1.36	1.26	0.47	0.93	0.26	0.21	0.67
i	0.15	1.02	0.06	3.00	0.26	6.81	0.06	3.96		2.40	2.24	16.00	0.90	0.51	0.17	0.08	0.00	0.20	0.63	1.57	1.14	0.27	0.08	0.27	0.07	0.47
j	0.05	0.19	0.00	0.06	0.00	0.55	1.43	0.31	2.07		0.45	0.47	0.06	0.00	0.00	0.13	0.00	0.00	0.06	0.06	0.07	0.00	0.06	0.18	0.39	0.00
k	0.47	2.80	0.13	3.91	0.05	5.39	0.48	5.59	1.02	0.64		11.90	0.40	0.32	0.10	0.57	0.31	0.21	1.33	1.24	0.30	0.31	1.23	4.11	0.23	0.97
l	0.00	1.23	0.00	2.42	0.00	7.93	0.00	7.14	1.94	1.13	3.78		0.43	0.30	0.00	0.13	0.00	0.06	0.43	0.51	1.17	0.13	0.06	0.28	0.14	0.46
m	0.80	0.27	0.00	1.33	0.00	3.01	0.30	3.70	1.15	0.45	0.87	3.15		5.44	0.19	0.88	0.93	3.24	0.58	0.45	3.09	0.33	4.09	0.35	0.40	0.49
n	1.55	2.29	0.52	3.18	0.35	4.60	0.90	17.00	0.86	0.65	1.59	5.30	18.40		2.42	3.34	1.59	5.11	1.51	1.12	4.95	1.10	2.69	1.07	0.61	1.50
o	4.22	12.50	10.10	9.66	0.99	0.89	6.15	1.01	0.31	0.65	0.61	0.21	0.40	0.78		7.77	3.73	0.96	2.74	0.38	2.09	1.60	0.95	0.51	0.45	1.04
p	0.62	0.76	0.07	0.47	0.11	0.35	7.35	0.87	0.34	0.43	0.30	0.38	0.37	0.13	1.04		21.80	0.33	0.92	0.06	1.06	0.07	0.19	0.25	0.77	0.13
q	0.65	1.04	0.34	2.17	0.19	0.80	9.42	0.32	0.26	0.59	0.40	0.34	0.07	0.20	2.34	1.41		0.82	0.71	0.31	1.46	0.33	0.39	0.21	1.06	0.35
r	1.52	2.40	2.17	3.70	1.00	12.70	0.23	8.12	2.51	0.89	2.85	4.77	9.38	3.95	1.63	1.41	1.06		2.78	1.56	4.57	1.37	2.04	2.61	1.13	3.65
s	9.58	0.64	0.41	1.06	7.85	1.02	0.49	0.42	0.15	0.07	1.04	0.63	0.37	0.55	0.43	0.70	0.15	0.41		0.20	0.22	0.06	0.72	1.37	0.35	4.28
t	0.36	5.30	0.56	5.98	0.15	9.99	0.50	4.09	2.66	0.86	4.65	7.17	0.34	0.68	0.68	0.27	0.21	1.95	1.19		6.06	1.06	0.51	0.59	0.36	0.87
u	0.72	3.01	0.32	4.39	0.34	3.44	0.73	6.01	2.63	1.99	2.41	5.04	1.21	3.08	1.86	1.66	1.36	2.35	1.52	3.02		2.46	2.87	1.38	0.47	0.77
v	0.34	0.33	0.44	0.25	0.08	1.11	0.06	0.28	0.09	0.56	1.62	0.77	0.72	0.44	0.25	0.24	0.24	0.56	0.82	0.78	1.88		26.40	3.96	10.40	1.88
w	0.33	0.39	0.00	0.31	0.25	1.26	0.00	0.66	0.13	0.51	3.12	0.52	0.66	0.35	0.11	0.13	0.12	0.39	1.04	0.79	0.51	3.38		1.95	4.71	0.80
x	1.21	0.51	0.33	0.75	0.20	1.27	0.24	0.60	0.21	0.23	8.99	0.84	1.21	0.51	0.13	0.32	0.08	0.42	3.38	0.47	0.27	1.02	2.87		1.45	5.63
y	0.00	0.07	0.00	0.00	0.00	0.42	1.76	0.13	0.12	0.53	1.23	0.28	0.07	0.00	0.00	1.77	1.11	0.05	0.12	0.36	0.43	2.14	2.37	2.81		2.04
z	1.22	0.12	0.40	0.26	0.51	0.89	0.16	0.35	0.14	0.17	2.86	0.90	0.49	0.62	0.22	0.05	0.06	0.92	2.46	0.28	0.38	0.94	1.69	8.12	0.81	

Stimulus

Response