Categorical membership modulates crowding: Evidence from characters

Josephine Reuther School of Psychology, University of Aberdeen, Aberdeen, UK

Ramakrishna Chakravarthi School of Psychology, University of Aberdeen, Aberdeen, UK

Visual crowding is generally thought to affect recognition mostly or only at the level of feature combination. Calling this assertion into question, recent studies have shown that if a target object and its flankers belong to different categories crowding is weaker than if they belong to the same category. Nevertheless, these results can be explained in terms of featural differences between categories. The current study tests if category-level (i.e., high-level) interference in crowding occurs when featural differences are controlled for. First, replicating previous results, we found lower critical spacing for targets and flankers belonging to different categories. Second, we observed the same, albeit weaker, category-specific effect when objects in both categories had the exact same feature set, suggesting that category-specific effects persist even when featural differences are fully controlled for. Third, we manipulated the semantic content of the flankers while keeping their feature set constant, by using upright or rotated objects, and found that meaning modulated crowding. An exclusively feature-based account of crowding would predict no differences due to such changes in meaning. We conclude that crowding results from not only the well-documented feature-level interactions but also additional interactions at a level where objects are grouped by meaning.

Introduction

A central question in vision research is to determine the mechanisms that underlie object recognition. Although object recognition appears effortless to us, it is an incredibly complex task that has not yet been fully understood (DiCarlo, Zoccolan, & Rust, 2012; Marr, 1982). One approach to studying object recognition is to examine a situation where it breaks down. Examining such a disruption allows reconstruction of the mechanisms underlying normal object recognition. A reliable and powerful example of such a breakdown is visual crowding. Crowding describes the reduced ability to resolve and recognize visual objects when surrounded by other stimuli (Bouma, 1970; Stuart & Burian, 1962). It is not an acuity limitation; even when visual acuity is sufficient to resolve a single object in the periphery, adding adjacent objects makes it more difficult to resolve that same object (Anstis, 1974; Pelli, Palomares, & Majaj, 2004).

Crowding studies have suggested that object recognition occurs over two stages. In the first stage features are detected. Pelli, Burns, Farell, and Moore-Page (2006) showed that recognition of even highly familiar objects like letters is contingent upon independent feature detection. In the second stage, the detected features are combined to form a representation that is then identified. However, features of nearby objects, if present, are also inappropriately combined with those of the target (Levi, 2008), leading to crowding. For example, Põder and Wagemans (2007) explicitly studied the mechanisms of the feature integration stage using gabors with multiple features (spatial frequency, color, and orientation), and discovered that participants' errors were not random, but driven by flanker properties. Furthermore, substitution errors were independently observed for all features of the multidimensional objects. They, therefore, argued that faulty feature integration causes crowding. In contrast, using cross-like stimuli with two feature dimensions (line orientation and position of bisection), Greenwood, Bex, and Dakin (2012) found that crowding appeared in an all-or-none fashion affecting already bound features. That is, incorrect feature integration took place for each feature, but occurred with the same probability for both object features. Thus, it remains unclear whether inappropriate feature binding in crowding takes place at the independent feature level

Citation: Reuther, J., & Chakravarthi, R. (2014). Categorical membership modulates crowding: Evidence from characters. *Journal of Vision*, *14*(6):5, 1–13, http://www.journalofvision.org/content/14/6/5, doi:10.1167/14.6.5.

doi: 10.1167/14.6.5 Received April 30, 2014; published October 16, 2014

(Põder & Wagemans, 2007) or at the object level (Greenwood et al., 2012). It has also been argued that the (mis)integration of features observed in crowding is prone to object mislocation and feature substitution (Hanus & Vul, 2013; Nandy & Tjan, 2007; Põder & Wagemans, 2007; Strasburger, 2005). However, the explanatory principle common to all these studies is that crowding is a bottom-up process where features are inappropriately bound (either by pooling or by feature/object substitution).

The maximal target-flanker distance at which crowding persists is proportional to the eccentricity of the peripheral object of interest (Bouma, 1970; Pelli & Tillman, 2008; Toet & Levi, 1992). This range is best described by Bouma's law (Bouma, 1970; Pelli & Tillman, 2008), which states that flankers interfere with the target's identification only if they are located within a region around the target spanning half the target eccentricity. Recently, Pelli and Tillman (2008) described crowding to be the same for all objects, arguing that the bottom-up process of faulty feature integration underlying crowding is agnostic about object identity or meaning. Evidence supporting this assertion comes from a study by Wallace and Tjan (2011), who directly compared the range of crowding for objects like lamps and rubber ducks and found that it was the same as that for letters, a stimulus set commonly used in crowding (e.g., Bouma, 1970; Pelli et al., 2004).

However, recent findings have begun to cast doubt on this simple picture of crowding. First, Martelli, Majaj, and Pelli (2005) showed that there is no crowding *within* nameable object parts, such as letters (part of a word) or eyes (part of a face), concluding that the letters of a word or the eyes of a face are processed holistically. However, they found crowding between letters and between face parts, which suggest that crowding acts not on a feature but on a letter or face-part level. Similarly, as mentioned above, Greenwood, Bex, and Dakin (2012) showed that pooling occurs over bound features rather than for each feature independently, suggesting that crowding occurs at the object level rather than at the feature level. Second, several studies have established that grouping between target and flankers strongly influences crowding (Chakravarthi & Pelli, 2011; Chicherov, Plomp, & Herzog, 2014; Livne & Sagi, 2007; Manassi, Sayim, & Herzog, 2012, 2013; Saarela, Westheimer, & Herzog, 2010). These results indicate that Gestalt principles that bind features and objects together, and not just low-level feature properties, play a crucial role in crowding. Third, work from the Whitney lab (Farzin, Rivera, & Whitney, 2009; Louie, Bressler, & Whitney, 2007) showed that crowding occurred even between high-level holistic representations by showing that an upright face target was more severely impaired by upright face flankers than by inverted ones, whereas no such effect of flanker orientation was found when the target face was inverted. Fourth, comparing crowding of letters and symbols, stronger crowding effects were found when only the flankers or both target and flankers were symbols compared to when they were letters. These studies either controlled for similarity in complexity between letters and symbols or used mirror images of the used letters as symbols. Since featural differences could be largely excluded, differences in familiarity were proposed to account for the found differences in crowding (Grainger, Tydgat, & Isselé, 2010; Huckauf, Heller, & Nazir, 1999); that is, familiar objects lead to less crowding than unfamiliar objects. As noted by Whitney and Levi (2011), taking these various findings of crowding studies together makes it hard to argue for a single locus of crowding. Indeed, imaging studies support the idea that crowding might be occurring at multiple stages of the visual processing pathway (Anderson, Dakin, Schwarzkopf, Rees, & Greenwood, 2012; Freeman, Donner, & Heeger, 2011). Crowding could therefore affect recognition of features, object parts, and/or whole objects.

Interestingly, the strength of crowding is modulated by the relation between target and flanker properties, such as their similarity in shape, contrast polarity or depth (Kooi, Toet, Tripathy, & Levi, 1994), color (Kennedy & Whitaker, 2010; Põder, 2007), and spatial frequency (Chung, Levi, & Legge, 2001). For example, Nazir (1992) manipulated the similarity between target and flankers by using a Landolt C target flanked by either four bars, Snellen Es, or Snellen Os. She found that the greater the similarity of shape between flankers and target, the worse the response accuracy for the orientation of the Landolt C target was. The general conclusion, based on studies using a large variety of objects, is that the more similar the objects are to each other, the more they crowd each other.

If crowding occurs at several stages of the visual processing pathway, then similarity between higher-level properties must also modulate crowding. That is, extrapolating the effect of similarity to the object level, crowding might be modulated by object properties such as meaning or categorical affiliation. Louie et al. (2007) and Farzin et al. (2009) provide indirect evidence for such similarity-based effects at the holistic representation level. However, this could be due to specialized processing just for upright faces and need not necessarily indicate high-level similarity effects. Direct evidence for category- or meaning-based similarity effects comes from a study conducted by Huckauf et al. (1999). They examined identification of letters flanked by upright letters, rotated letters, numbers, and letter-like nonmeaningful stimuli. They discovered that crowding was stronger for nonmeaningful flankers than for meaningful flankers; further, within the category of meaningful flankers, crowding was stronger when the flankers belonged to the same category as the target than if they belonged to different categories. Thus, two factors, object familiarity and category, were argued to play a role in crowding. First, differences in crowding caused by letter, rotated letter, and letter-like meaningless flankers might be based on the differences in exposure to these objects, or in other words, their familiarity or meaning, with familiar flanker objects somehow interfering less with target recognition. Second, the difference between the effects of letter and number flankers on letter crowding might be interpreted as the result of differences in their categorical affiliation, with less crowding of letters by number flankers (a category-specific crowding effect based on similarity at the object level). Note that the latter finding cannot be explained by possible familiarity differences between letters and numbers. Since unfamiliar flankers lead to more crowding than familiar ones, in order to explain this result, letters would have to be less familiar than numbers. However, the opposite is the case, with letter being more familiar than numbers.

However, to conclude that object category membership or meaning explains the findings of Huckauf et al. (1999) might be premature. Lower crowding for a letter target surrounded by number flankers when compared to letter flankers might be sufficiently described by lower between-group than within-group feature similarity; in other words, letters just share more features with other letters than they do with numbers, leading to the observed differences in crowding. Furthermore, even if category-membership has an effect on crowding, it remains unclear if this is a general effect applicable to a variety of categories or if it is specific to letters as targets.

Proceeding from these findings we wanted to test if there is a genuine category-specific effect (similar category objects crowd more than dissimilar category objects) and whether this category-specific effect could be generalized beyond letter targets. That is, we sought to replicate Huckauf et al. (1999) and test if their results were merely due to low-level local feature differences between categories or due to categorical (dis)similarity. Accordingly, we tested if letters crowded letters more than numbers crowded letters and vice versa (Experiment 1). Next, we also designed a new font wherein target objects of both categories had the exact same local feature content (but differed in their meaning). Crowding such objects would indicate whether featural differences drove the observed category-specific effects (Experiment 2). Finally, we tested if crowding was modulated by flanker meaning. We examined if crowding of a number target by meaningful and meaningless flankers would be the same when they have the exact same features (Experiment 3).

General methods

We will first describe the methodology common to all experiments and then describe experiment specific details separately.

Observers

Forty-seven observers aged 17 to 49 years took part in this study. All had normal or corrected-to-normal vision. Twenty-one participated in Experiment 1, 21 in Experiment 2, and 15 in Experiment 3. All participants gave written, informed consent. The study was designed and conducted under the approval of the Psychology Ethics Committee at the University of Aberdeen.

Material and stimuli

Stimuli were generated in MATLAB using Psychophysics toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) and displayed on a 19-in. Sony CRT screen (Sony Trinitron GDM-F520, Sony Corporation, Tokyo, Japan) with a frame rate of 100 Hz and resolution of 1024 × 728 pixels. Viewing distance was set to 57 cm secured by the use of a chinrest.

Letters and numbers served as targets and flankers in a 2 (target type: letters or numbers) by 2 (flanker type: letters or numbers) study design, leading to four conditions. Stimuli were black ($L_v = 0.25$ cd/m^2) on white background ($L_v = 91.5$ cd/m^2) with a high contrast ($C_{michelson} = 0.99$). The target was presented 10° from fixation on the horizontal meridian randomly either in the left or the right visual field. It was flanked by four flankers, one in each cardinal direction (left, right, top, and bottom; see Figure 1). The center-to-center spacing between the target and its flankers was controlled by the QUEST algorithm (Watson & Pelli, 1983). This algorithm was used to estimate the critical spacing at which target identification performance would be 82%.

Procedure

Each experiment consisted of several conditions (four in Experiments 1 and 2; three in Experiment 3). Conditions were blocked. Up to five thresholds (three in Experiment 1) were estimated for each condition and then averaged. Each block consisted of 32 trials.

At the beginning of each block the observer was informed about the category of the target, but the flanker category was not mentioned. Each block was started with a key press. A fixation mark appeared at

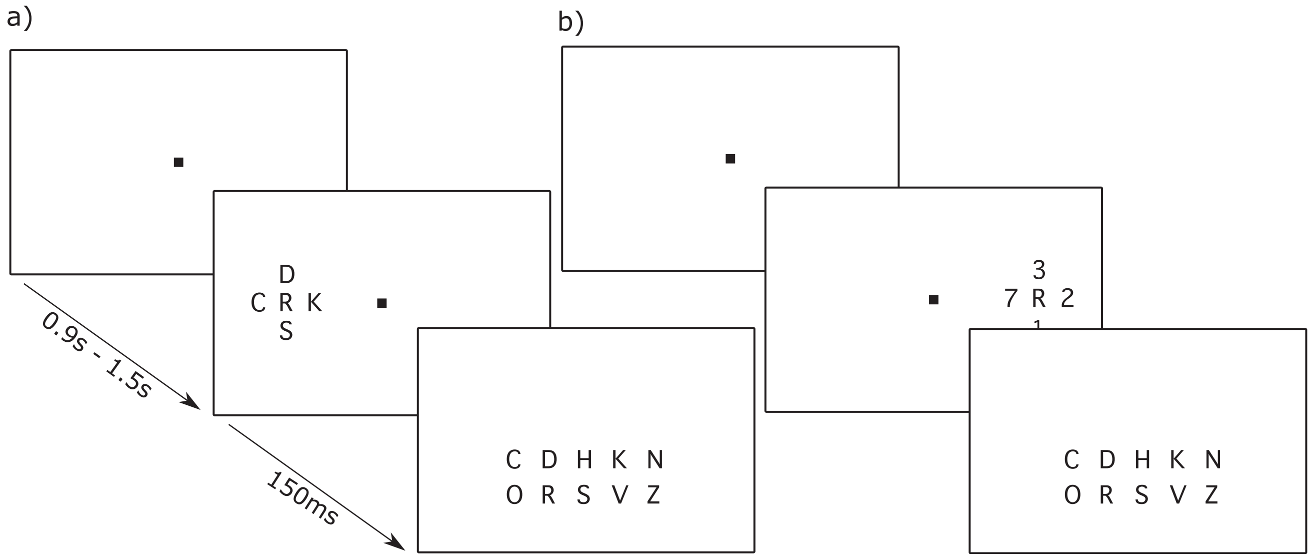

Figure 1. General procedure: A target with 4 flankers is displayed randomly in the right or left visual field for 150 ms after varying onset times. a) The left series shows the condition where target and flankers belong to the same category; b) The series on the right displays characters that differ in category; both situations apply to number targets as well.

the center of the screen and stayed on throughout the block. Target and flankers were randomly displayed in the left or right visual field after varying onset times (0.9–1.5 s after trial onset) for 150 ms, ensuring that no eye movements could be made. After the stimuli disappeared, all possible target stimuli were displayed on the screen for mouse response. The next trial started automatically after the observer's response. Auditory feedback was provided.

Experiment 1: Crowding of letters and numbers

Experiment 1 was aimed at replicating and extending the results of Huckauf et al. (1999). They examined the effect of letter and number flankers on letter targets to test the effect of category affiliation. We extended that approach with a two (target: letter, number) by two (flanker: letters, numbers) factorial design. This would probe if the effect they observed, that number flankers crowd letter targets less than letter flankers do (that is, same category flankers are more effective crowders than different category flankers), can be replicated and generalized to numbers as targets. It could be that their results were specific to letters, since (a) in general, we are more familiar with letters and this could have helped process the letter flankers more, somehow leading to more crowding, and (b) there might be featural differences between letter and number flankers that could have reduced crowding between them, since we know that flankers dissimilar to the target result in less crowding than similar flankers (Kooi et al., 1994; Nazir, 1992). Weaker crowding when target and flankers belong to different categories, *irrespective of the particular target category* (letter or number), compared to when target and flankers belong to the same category, would suggest that crowding also occurs at a higher level of processing.

Stimuli and procedure

Ten uppercase letters (C, D, H, K, N, O, R, S, V, and Z) and 10 numbers (from 0 to 9) served as targets and flankers. All characters were presented with 44 pt Geneva font. Letters subtended, on average, 2.10° of visual angle, vertically. There were four conditions: (a) letter targets flanked by letters, (b) letter targets flanked by numbers, (c) number targets flanked by letters, and (d) number targets flanked by numbers. Experiment 1 had a within-subject design, and we estimated critical spacing in all four target-flanker conditions for each observer.

The experiment started with two practice blocks, one per target category (letters, numbers), followed by 12 experimental blocks, three for each target-flanker combination.

Results and discussion

We tested if target-flanker category membership had any effect on crowding even when two different target categories were used. A 2 (target: letter or number) by 2 (flanker: letters or numbers) repeated-measures ANOVA

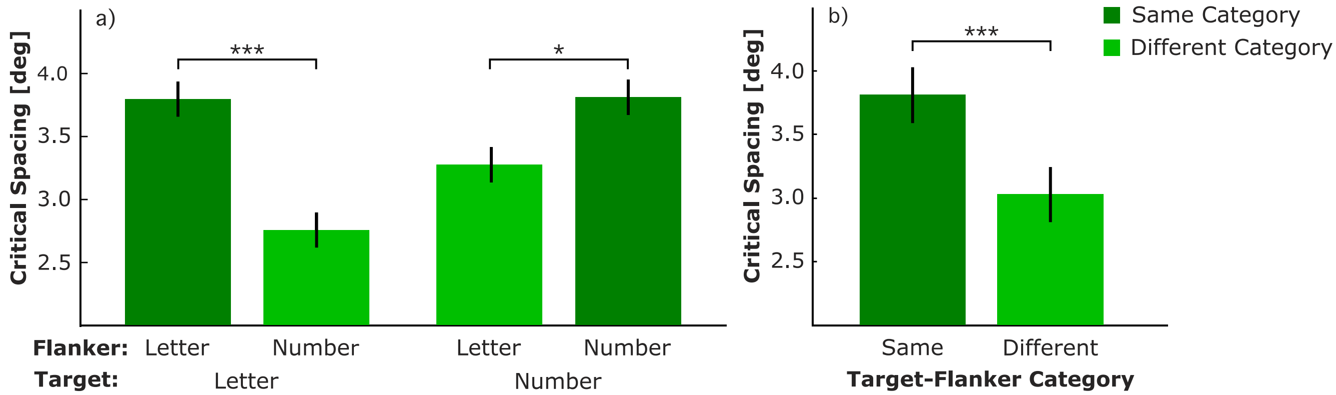

Figure 2. Results of Experiment 1: Critical spacing is modulated by the target-flanker category similarity, with flankers of the same category as the target causing more crowding than flankers of a different category. This was irrespective of the specific category tested. Error bars are ± 1 $SEM_{L\&M}$ (within-subject error bars).

was performed to examine the influence of stimulus category on the amount of crowding. Follow-up pairwise *t* tests (Bonferroni-corrected) were conducted when necessary. Four observers were excluded from the study for not following instructions, since the observed thresholds were not reliable (the standard deviation of the threshold estimates was greater than 0.2 log units). The results for all four combinations of target and flanker category are plotted in Figure 2.

Crowding is shown to be independent of target type, $F(1, 64) = 0.89$, $p = 0.349$, $\eta_p^2 = 0.014$, and flanker type, $F(1, 64) = 0.81$, $p = 0.372$, $\eta_p^2 = 0.012$. However, there was a strong interaction indicating that categorical affiliation affects crowding, $F(1, 64) = 7.68$, $p = 0.007$, $\eta_p^2 = 0.107$. Pairwise *t* tests showed that crowding is lower for conditions where a letter target is flanked by a number flanker than when it is flanked by letters, $t(16) = 6.68$, $p < 0.001$. Similarly, when a number is flanked by letters, crowding is weaker than when it is flanked by numbers, $t(16) = 2.92$, $p = 0.020$. Collapsed across categories (Figure 2b), same-category crowding is stronger than opposite-category crowding, $t(16) = 6.24$, $p < 0.001$.

Experiment 1 replicated Huckauf et al.'s (1999) findings that letters are less crowded by numbers than by letters and extended their findings to numbers as targets, showing that the effect is generalizable to other categories. These results suggest that crowding not only occurs at the feature integration stage, but also at the level where objects are categorized.

However, this experiment does not completely rule out the possibility that featural differences can account for category-specific effects. If we assume that there are fewer featural differences within a category than between different categories, then we would expect less crowding between objects belonging to different categories than for objects belonging to the same category, independent of category type. For example, numbers share fewer features with letters than with other numbers, leading to weaker crowding between number targets and letter flankers than for number target and number flankers. Experiment 2 was designed to test this alternative low-level, feature-based account.

Experiment 2: Controlling for featural differences

In Experiment 2, we controlled for featural differences between two categories to test if these differences accounted for the findings of Experiment 1. We designed a specific font similar to the digital-seven and calculator fonts and used a restricted set of characters that eliminated any local featural differences between the objects belonging to two separate categories. Any observed category-specific crowding effects can therefore not be explained by the presence of featural differences.

Stimuli and Procedure

The font designed for Experiment 2 consisted of only seven straight lines in the style of a calculator font. The individual letters subtended 2.10° of visual angle, just as in Experiment 1. The letters I, J, O, and P, and the numbers 0, 1, 7, and 9 served as targets. As can be seen in Figure 3, the targets I and 1 and O and 0 are identical and the other targets, 7 and J and 9 and P, are mirror images of each other. Thus, the characters used as targets all had the exact same feature sets, irrespective of the category they belonged to. We used the letters: A, E, J, P, U, and Y and numbers: 2, 3, 4, 6, 7, and 9 as flankers.

Only characters that were clearly perceived as belonging to one of the two categories were used as

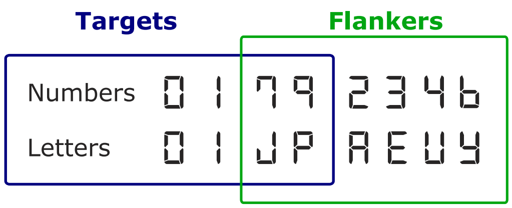

Figure 3. Targets and flankers used in Experiment 2: Target letters and numbers are identical or mirror images of each other, whereas flankers are chosen to clearly represent a character corresponding to the respective category.

flankers. That is, for example, since a digital 5 can be perceived as an S, it was excluded from the flanker set. Three of the used flanker characters were mirror images of each other (7/J, 9/P, and 3/E). Among the rest, two of the number characters had one stroke less than the corresponding letter characters (2/A and 4/Y). To ensure that a given observer would perceive the targets as belonging to their assigned category (letters or numbers), target type was a between-subject factor. Therefore, each observer participated in only two conditions: For example, number targets flanked by numbers or by letters (a different participant would be tested with letter targets flanked by numbers or letters). Therefore, in distinction to Experiment 1, Experiment 2 had a mixed design. As in Experiment 1 the observers performed two practice blocks, this time followed by five blocks per flanker condition.

Results of Experiment 2

A mixed ANOVA was conducted to analyze the influence of stimulus category on the amount of crowding. Here, target type served as between subject factor and flanker type as within subject factor. One observer was excluded from the statistics for noncompliance, using the same criterion as before. The results of Experiment 2 are plotted in Figure 4.

As in Experiment 1 no effects of target type, $F(1, 18) = 2.199$, $p = 0.155$, $\eta_p^2 = 0.109$, and flanker type, $F(1, 18) = 0.004$, $p = 0.949$, $\eta_p^2 < 0.001$, were observed. However, once again an interaction was found, $F(1, 18) = 5.090$, $p = 0.037$, $\eta_p^2 = 0.220$. Collapsing the results across numbers (category-specific effect for number targets: $t(9) = 1.647$, $p = 0.134$) and letters (category-specific effect for number targets: $t9 = 1.554$, $p = 0.115$) in order to compare same-category thresholds with different category thresholds (Figure 4b), irrespective of target type, we found that crowding was lower for conditions where targets and flankers differed in category affiliation, $t(19) = 2.318$, $p = 0.032$.

We noted earlier that some letter flankers had additional strokes relative to number flankers. Could this difference in complexity (as measured by number of strokes) explain these results? It is unclear what the effect of the differences in complexity between the flanker types would be on crowding. It could be that the similarity in complexity between number flankers and both target types (since all of these characters are less complex than letter flankers) increases the crowding induced by number flankers (Zhang, Zhang, Xue, Liu, & Yu, 2009) or that the higher number of features in the letter flankers allows them to be bound with the target features leading to more crowding by the letter flankers (similar to increasing the number of flankers; Põder & Wagemans, 2007). Either way, this predicts only a main effect of flanker type since both target types should be affected similarly (as they share the same level of complexity, in terms of stroke number). In general, any differences in the feature sets of the flankers only predicts a main effect of flanker type and not an interaction, as the category-specific effect would predict. As described above, no such effect of flanker type, $F(1, 18) = 0.004$, $p = 0.949$, $\eta_p^2 < 0.001$, was observed, ruling out an effect based on the possible (if small) complexity differences between the two flanker types.

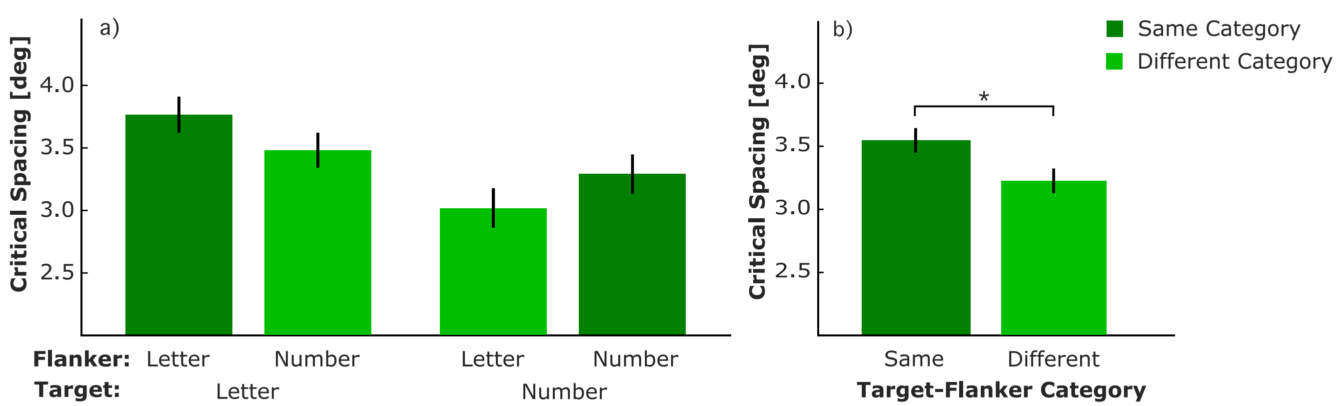

Figure 4. Results of Experiment 2: Critical spacing is dependent on similarity of categorical affiliation of target and flanker. Crowding is weaker when target and flankers differ in categorical affiliation. Error bars are ± 1 $SEM_{L\&M}$ (within subject error bars).

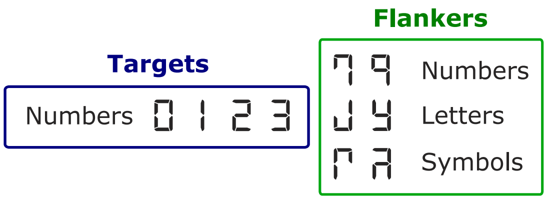

Figure 5. Targets and flankers used in Experiment 3.

Comparing Figures 2 and 4, it can be noticed that the category-specific effect (stronger crowding for within category objects than for between category objects) is weaker here. To quantify this observation, we compared the strength of the category-specific effect in the two experiments. We conducted a mixed model ANOVA with Experiments 1 or 2 as the between subject factor and target-flanker category (same or different) as the within subject factor. Results showed a main effect of target-flanker category, $F(1, 52) = 31.976$, $p < 0.001$, $\eta_p^2 = 0.381$, confirming the general conclusion about the existence of a category-specific effect. Although there was no main effect of experiment, $F(1, 52) = 0.018$, $p = 0.893$, $\eta_p^2 < 0.001$, there was a significant interaction between experiment and target-flanker category, $F(1, 52) = 6.159$, $p = 0.016$, $\eta_p^2 = 0.106$. Posthoc t tests revealed that the category-specific effect (difference between same and different target-flanker category thresholds) was weaker in Experiment 2 than in Experiment 1, $t(46) = 1.647$, $p = 0.006$. That is, the difference between same and different category critical spacing is smaller here. This could be due to several reasons: the use of a specific font, target discriminability (perhaps easier identification task in Experiment 2), or absence of featural differences (that is, featural differences partially contributed to the larger effect observed in Experiment 1). This experiment cannot distinguish between these alternatives. However, we believe that it cannot be the font since it has been shown that font (unless it is complex) does not affect crowding (Pelli et al., 2004). Further, even though the number of response alternatives might not play a role here, the discriminability between targets might. Highly discriminable targets (e.g., two gratings differing in orientation by 90°) are less susceptible to crowding than very similar ones (differing by 10°; Pelli et al., 2004). However, the targets we used were not that different from each other in Experiment 2 (sometimes differing by just one line). Further, using highly discriminable letters as targets (such as x and o) still produces considerable crowding (e.g., Nandy & Tjan, 2007). Hence, it is likely that the difference between Experiments 1 and 2 is due to the difference in feature similarity (or lack of it in Experiment 2) between categories.

In summary, even when the target characters of different categories have the exact same feature set, critical spacing is smaller when target and flankers differ in category, suggesting that at least a part of the category-specific effect on crowding is due to categorical differences and not due to featural differences. These findings underpin the role of higher order processes in crowding.

Experiment 3: Crowding by meaningless stimuli

To further test if higher-level processes were involved in crowding, we examined if flanker meaning influences critical spacing. If a feature-based account were sufficient to explain the findings of the prior experiments, it would predict that changing flanker orientation, which does not change low-level local properties, should have no influence on critical spacing. That is, although meaning and therefore categorizability are lost by inverting a character, since features remain the same, critical spacing should not be affected. Previously, Huckauf et al. (1999) showed that for letters as targets, meaningless flankers caused stronger crowding effects than meaningful flankers. A second goal of this experiment was to replicate these findings with number targets, which would enable generalizability of the effect of meaning in crowding, while crucially keeping the feature set constant across various categories. If the semantic meaning of an object modulates crowding, it would imply that crowding cannot merely be a low-level feature integration process; higher-level processes that have access to the meaning of the flankers must be influencing the pooling of information that underlies crowding.

Stimuli and procedure

The same font as in Experiment 2 was used. The targets were always numbers, 0–3. The letters J and Y and the numbers 7 and 9 served as flankers, since they shared the same features (in the designed font). Note that 9 and Y did not share the exact same feature set, but had the same number of vertical and horizontal elements, with one horizontal element displaced in position. Meaningless flankers (symbols) were rotations and/or mirror images of the used letter flankers. Targets and flankers can be seen in Figure 5.

Unlike in Experiments 1 and 2, the target was displayed at 15° eccentricity, with a viewing distance of 38 cm. This eccentricity was chosen to avoid floor effects found in a pilot experiment (where the same targets were used with only the J flanker in three orientations: J, 7, and a meaningless mirror image of 7). Each object subtended 2.10° of visual angle.

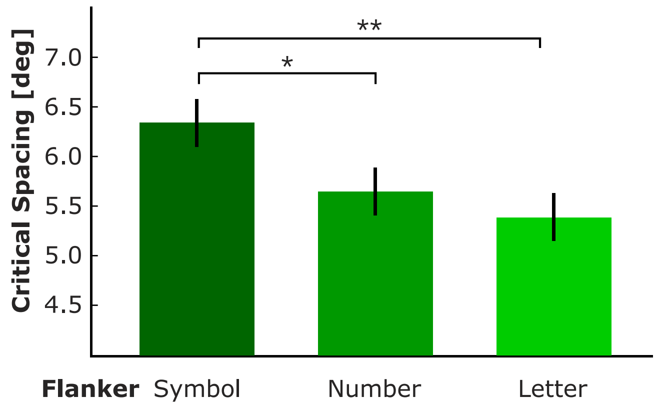

Figure 6. Results of Experiment 3: Critical spacing for a number target was found to be modulated by meaning, with meaningless flankers leading to stronger crowding than meaningful. Error bars are ± 1 $SEM_{L\&M}$ (within-subject error bars).

The observers performed three practice blocks, one for each flanker condition followed by five blocks per flanker condition.

Results

Experiment 3 tested if meaning modulated crowding by comparing the effects of meaningful flankers (letters and numbers) with that of meaningless flankers (symbols). A one-way repeated-measures ANOVA with flanker type (number, letter or symbol) as a factor was performed. The results of Experiment 3 can be seen in Figure 6.

One participant was excluded from the statistics for noncompliance since s/he had an obvious decline of performance over time independently of the flanker condition. Crowding was modulated by flanker meaning, $F(1, 12) = 2.199$, $p = 0.011$, $\eta_p^2 = 0.531$, with meaningless symbol flankers leading to stronger crowding than numbers, $t(13) = 2.276$, $p = 0.04$, and letters, $t(13) = 6.68$, $p = 0.006$. No other effects were significant. These results show that meaningless flankers cause more crowding than meaningful ones, even though the features in meaningful and meaningless flankers were identical (to at least the letter flankers). This strengthens the claim that crowding affects object recognition not only at a feature integration level but also at or after a level were meaning is processed.

It has been shown earlier that crowding caused by meaningless flankers exceeds that by meaningful flankers (Huckauf et al., 1999), a finding replicated here. However, the conclusion drawn from Experiments 1 and 2, that crowding is strongest when targets and flankers share categorical affiliation, would have predicted the contrary, as the targets do not share category membership with meaningless flankers. Therefore, crowding for number targets in Experiment 3 should have been strongest when surrounded by number flankers. Furthermore, crowding induced by letter and symbol flankers should have been similar, especially since the feature sets of those two categories were identical. Our findings contradicted this prediction, but confirmed and extended the meaning (or familiarity) dependent modulation of crowding previously observed by Huckauf et al. (1999). This suggests that the underlying mechanisms for meaning- and category-dependent modulation of crowding might be different. Nevertheless, it indicates that higher-level object properties (meaning or familiarity and category membership) modulate crowding, implying that crowding cannot merely be faulty feature integration. It further demonstrates that the meaning dependent modulation is not specific to letters but is generalizable to other objects. We discuss the implications of these findings, specifically the modulation of crowding by meaning and how it might differ from category-dependent modulation, in the Discussion section.

An alternative explanation could be that although the features were the same, rotating the objects might change the relative positions of the features, which might affect the number of flanker features that abut the target. Such positional differences of features would then modulate crowding with stronger crowding caused by flankers that had the right feature locations (say, the meaningless flankers here). This is unlikely, since flankers were presented in all four locations and thus all parts of the flankers abutted the target on different trials. This would be true for all flanker types and hence the target flanker interactions should be similar across category as long as the local features are the same, as they are in our stimulus set. That is, the same features are exposed to the target for all flanker types.

Unlike in Experiments 1 and 2, we did not find a category-specific crowding effect (a difference in crowding caused by number and letter flankers; $t(13) = 0.630$, $p = 0.539$). A possible reason for not finding an effect of category-membership in this experiment might be due to the higher eccentricity (15°) tested here compared to the other experiments (10°). Huckauf et al. (1999) found that the effect of category-membership was strong at 1°, weaker at 4°, and weakest at 7°. That is, the effect for category-membership reduces considerably with eccentricity. It is possible that by 15° (as tested here), there is no category-specific effect, although there is a trend that appears in the expected direction. Furthermore, it is known that central retinal areas are better at meaning extraction than peripheral ones (Latfiam & Whitaker, 1996). Hence, it is possible that the far periphery is unable to disambiguate category membership easily, which might lead to the lack of category-specific effect observed here. Further testing needs to be done to determine if this is the case.

Discussion

This study tested the hypothesis that crowding can be explained in terms of purely low-level feature interactions. If, on the other hand, crowding occurs at multiple levels of the visual processing hierarchy, then category membership should affect crowding, even when the feature set is kept constant. Experiment 1 replicated and extended a category-specific similarity effect documented by Huckauf et al. (1999), where objects belonging to the same category crowd each other more than objects belonging to different categories. Experiment 2 showed that controlling for feature similarity does not eliminate this effect. The third experiment indicated, contrary to feature-based accounts, that crowding was modulated by meaning (present vs. absent), with meaningless stimuli being more potent at inducing crowding than meaningful stimuli, even when the feature set was the same in all conditions. These findings support a growing consensus that crowding occurs at multiple levels (Anderson et al., 2012; Farzin et al., 2009; Freeman et al., 2011; Greenwood et al., 2012; Louie et al., 2007; Manassi et al., 2012, 2013; Whitney & Levi, 2011).

Previously, Huckauf et al. (1999) had shown that letter targets were more strongly crowded by letter flankers than by number flankers, but most of all by meaningless flankers. We replicated those findings here, while ensuring that the features in all categories were the same, and extended the results to other categories, such as number targets, thus demonstrating the generalizability of this effect. It must be noted that Huckauf et al. (1999) used accuracy of identification as the dependent measure, whereas we directly measured critical spacing and found similar results. In other words, same-category flankers interact with the target from much larger distances than do different category flankers. Our results also show that the category-specific similarity effect could not be merely due to featural differences between the tested categories.

The finding that flanker orientation of featurally identical flankers has an effect on critical spacing in Experiment 3 strengthens the conclusion that an exclusively feature-based account is not sufficient to explain the converging findings in the crowding literature. Here, orientations that preserved meaning (whether letters or numbers) resulted in less crowding than those that did not preserve meaning (meaningless symbols). Additionally it has been shown that flanker orientation had an effect on crowding only for meaningful targets, such as letters (Huckauf et al., 1999) and faces (Farzin et al., 2009; Louie et al., 2007). When the target was inverted and therefore not perceived as a letter or a face (that is, when they were devoid of meaning), there was no effect of flanker orientation. An orientation or feature-based account for crowding would have predicted that an inverted target would be hampered more severely by inverted flankers. Showing that orientation plays a role only for upright targets leads to the conclusion that it is meaning and not orientation that is the driving factor. It might be that meaningful targets are perceived as a whole, as was proposed by Martelli et al. (2005), and unknown objects are perceived as a collection of the features they are made up of. If so, the features of recognizable objects might be tightly bound, whereas those of unknown objects are more loosely bound, which can then be inappropriately integrated with target features leading to stronger crowding. That is, stable representations of objects, such as those elicited by letters and numbers, lead to less crowding than unstable representations, such as those of meaningless symbols.

An alternative, but related, explanation might be one suggested by the finding by Grainger et al. (2010) that crowding between (known) symbols was stronger than that between letters. The authors concluded that the amount of exposure to objects would influence the size of the integration field for such objects, with familiar objects interacting only at short distances. There is evidence that the size and shape of the integration field can be modulated by experience (Chung, 2007; Nandy & Tjan, 2007). Extensive learning at a particular location can reduce the extent of crowding by almost 40% (Chung, 2007). Our results that meaningful flankers caused less crowding than meaningless ones are in line with the above findings that familiarity affects critical spacing.

The two proposed explanations are compatible with each other. It is possible that extended exposure to objects can lead to the formation of stable representations and hence more holistic processing of these objects. These representations will, therefore, only interact with other representations within a short distance. Indeed, it is known that highly familiar stimuli are processed rapidly and automatically (Li, VanRullen, Koch, & Perona, 2002; VanRullen, Pascual-Leone, & Battelli, 2008). Such stimuli might be processed through hard-wired pathways, which would reduce their susceptibility to crowding and their contribution to crowding relative to novel stimuli.

We combine these findings and ideas in a tentative proposal that might explain our results. There are hard-wired pathways that facilitate the processing and binding of highly familiar objects, whereas unfamiliar objects are processed effortfully and binding among their features is weak. This would account for the effect of flanker meaning on crowding, where the features of meaningless flankers can easily combine with those of the target. Only familiar objects containing meaning can later (after features are combined) and further interact, based on higher order similarities, which accounts for the findings that category membership and flanker meaning modulate crowding.

However, access to these hard-wired pathways becomes more difficult with more peripheral presentation (which is perhaps the underlying reason for the increasing difficulty in reading with eccentricity; see, for example, Chung, 2007; Latfiam & Whitaker, 1996), probably because these objects were most likely learned in or near the fovea. Hence it becomes harder to distinguish the category membership of objects at higher eccentricities. Therefore, category-specific effects are noticeable at near eccentricities but not at farther ones. This interpretation makes a testable prediction: the distinction between meaningful and meaningless flankers should disappear at large eccentricities (say, 30°), as at those eccentricities stimuli should have little to no access to hard-wired pathways. On the other hand, the meaningful-meaningless distinction might be independent of the determination of category membership, with the latter dependent on eccentricity, but not the former. Future experiments could help distinguish these alternatives.

Crowding and texture perception

Crowding mechanisms are thought to compute textures that pool information over multiple nearby objects (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). In crowding, there is growing evidence that information about individual objects is not completely lost but is pooled or averaged. This has been shown for orientation (Parkes et al., 2001), position (Greenwood, Bex, & Dakin, 2009), and even for emotion discrimination (Fischer & Whitney, 2011). Such pooling, averaging, or texturization can, therefore, represent a statistical summary of the objects present in the periphery (e.g., Balas, Nakano, & Rosenholtz, 2009; Freeman & Simoncelli, 2011). It is known that a statistical summary of objects, which is or can be computed automatically and preattentively (Alvarez & Oliva, 2009) and which does not require identification of individual objects (Ariely, 2001; Chong & Treisman, 2003), can modulate perception and guide attention and eye movements (Alvarez, 2011; Duncan & Humphreys, 1989; Oliva & Torralba, 2007; Wolfe, Oliva, Horowitz, Butcher, & Bompas, 2002). For example, Wolfe et al. (2002) argued that rapid segmentation (a product of appropriate texturisation) is a crucial first step in search and that the effectiveness of this segmentation affects the later, slower identification process. Further, averaging has been shown to reduce errors (Alvarez, 2011), which benefits perception and behavior guidance. Thus, the output of crowding mechanisms, in the form of a texture, can be utilized by the visual system, even though information about individual items cannot be retrieved. Our results show that higher-order attributes of objects, such as category membership, modulate crowding mechanisms, with same-category objects more likely to crowd each other. Taken together, these findings suggest that higher-level interference in crowding might enable objects belonging to the same category to be pooled and form a texture, and objects belonging to different categories to be less likely to be pooled. Conceptually similar findings have been reported by the Herzog lab, where objects belonging to the same perceptual group crowd more easily than objects belonging to different perceptual groups (Manassi et al., 2012, 2013; Saarela, Sayim, Westheimer, & Herzog, 2009; Saarela et al., 2010; see also Chakravarthi & Pelli, 2011; Livne & Sagi, 2007). Such preferential pooling, and hence texturization, of objects with the same category membership might allow a more efficient and precise representation of information, and conversely, allow outlier detection (or a more effective segmentation) of items that do not belong to the same category (Alvarez, 2011). Category-level interactions in crowding leading to texture formation of different effectiveness might therefore affect perception and behavior.

Conclusion

We found that meaning influences critical spacing with lower crowding caused by meaningful flankers. Furthermore, we showed that characters crowd less when target and flankers differ in category, independent of their categorical affiliation. This was true even when featural differences were controlled for. A feature-based account cannot fully account for such a category-specific effect. We suggest that higher-level processes modulate crowding through feedback to the feature integration stage.

Keywords: crowding, object category, meaning, object recognition, feature integration

Acknowledgments

We wish to thank Michael H. Herzog for suggesting that changing orientation should eliminate meaning and therefore category membership, which led to Experiment 3.

Commercial relationships: none.
Corresponding author: Josephine Reuther.
Email: j.reuther@abdn.ac.uk.

Address: School of Psychology, King's College, University of Aberdeen, UK.

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. [PubMed]

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, USA*, *106*(18), 7345–7350. [PubMed]

Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G., & Greenwood, J. A. (2012). The neural correlates of crowding-induced changes in appearance. *Current Biology*, *22*(13), 1199–1206. [PubMed]

Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14*(7), 589–592.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162. [PubMed]

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision*, *9*(12):13, 1–18, http://www.journalofvision.org/content/9/12/13, doi:10.1167/9.12.13. [PubMed] [Article]

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177–178.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. [PubMed]

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

Chicherov, V., Plomp, G., & Herzog, M. H. (2014). Neural correlates of visual crowding. *NeuroImage*, *93*(Part 1), 23–31. [PubMed]

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404.

Chung, S. T. L. (2007). Learning to identify crowded letters: Does it improve reading speed? *Vision Research*, *47*(25), 3150–3159. [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. [PubMed] [Article]

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How Does the Brain Solve Visual Object Recognition? *Neuron, 73*(3), 415–434. [PubMed] [Article]

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458.

Farzin, F., Rivera, S. M., & Whitney, D. (2009). Holistic crowding of Mooney faces. *Journal of Vision*, *9*(6):18, 1–15, http://www.journalofvision.org/content/9/6/18, doi:10.1167/9.6.18. [PubMed] [Article]

Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology*, *106*(3), 1389–1398. [PubMed] [Article]

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, http://www.journalofvision.org/content/11/4/15, doi:10.1167/11.4.15. [PubMed] [Article]

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*(9), 1195–1201. [PubMed] [Article]

Grainger, J., Tydgat, I., & Isselé, J. (2010). Crowding affects letters and symbols differently. *Journal of Experimental Psychology: Human Perception & Performance*, *36*(3), 673–688. [PubMed]

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. [PubMed] [Article]

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, http://www.journalofvision.org/content/12/3/18, doi:10.1167/12.3.18. [PubMed] [Article]

Hanus, D., & Vul, E. (2013). Quantifying error distributions in crowding. *Journal of Vision*, *13*(4):17, 1–27, http://www.journalofvision.org/content/13/4/17, doi:10.1167/13.4.17. [PubMed] [Article]

Huckauf, A., Heller, D., & Nazir, T. A. (1999). Lateral masking: Limitations of the feature interaction account. *Perception & Psychophysics*, *61*(1), 177–189. [PubMed]

Kennedy, G. J., & Whitaker, D. (2010). The chromatic selectivity of visual crowding. *Journal of Vision*, *10*(6):15, 1–13, http://www.journalofvision.org/content/10/6/15, doi:10.1167/10.6.15. [PubMed] [Article]

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

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. [PubMed]

Latfiam, K., & Whitaker, D. (1996). A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Research*, *36*(17), 2665–2674. [PubMed] [Article]

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*(5), 635–654. [PubMed] [Article]

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, USA*, *99*(14), 9596–9601. [PubMed] [Article]

Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision*, *7*(2):4, 1–12, http://www.journalofvision.org/content/7/2/4, doi:10.1167/7.2.4. [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, http://www.journalofvision.org/content/7/2/24, doi:10.1167/7.2.24. [PubMed] [Article]

Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *Journal of Vision*, *12*(10):13, 1–14, http://www.journalofvision.org/content/12/10/13, doi:10.1167/12.10.13. [PubMed] [Article]

Manassi, M., Sayim, B., & Herzog, M. H. (2013). When crowding of crowding leads to uncrowding. *Journal of Vision*, *13*(13):10, 1–10, http://www.journalofvision.org/content/13/13/10, doi:10.1167/13.13.10. [PubMed] [Article]

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, *5*(1):6, 58–70, http://www.journalofvision.org/content/5/1/6, 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, http://www.journalofvision.org/content/7/2/5, 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. [PubMed]

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527. [PubMed]

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. [PubMed]

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*(28), 4646–4674. [PubMed] [Article]

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*(10), 1129–1135. [PubMed]

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

Põder, E. (2007). Effect of colour pop-out on the recognition of letters in crowding conditions. *Psychological Research*, *71*(6), 641–645. [PubMed]

Põder, E., & Wagemans, J. (2007). Crowding with conjunctions of simple features. *Journal of Vision*, *7*(2):23, 1–12, http://www.journalofvision.org/content/7/2/23, doi:10.1167/7.2.23. [PubMed] [Article]

Saarela, T. P., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, *9*(2):5, 1–11, http://www.journalofvision.org/content/9/2/5, doi:10.1167/9.2.5. [PubMed] [Article]

Saarela, T. P., Westheimer, G., & Herzog, M. H. (2010). The effect of spacing regularity on visual crowding. *Journal of Vision*, *10*(10):17, 1–7, http://www.journalofvision.org/content/10/10/17, doi:10.1167/10.10.17. [PubMed] [Article]

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, *5*(11):8, 1024–1037, http://www.journalofvision.org/content/5/11/8, doi:10.1167/5.11.8. [PubMed] [Article]

Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty. Its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology*, *53*, 471–477.

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*(7), 1349–1357. [PubMed]

VanRullen, R., Pascual-Leone, A., & Battelli, L. (2008). The continuous wagon wheel illusion and the "when" pathway of the right parietal lobe: A repetitive transcranial magnetic stimulation study. *PLoS ONE*, *3*(8), e2911. [PubMed] [Article]

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

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. [PubMed] [Article]

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research*, *42*(28), 2985–3004. [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. [PubMed] [Article]