Published in final edited form as:

Atten Percept Psychophys. 2015 February ; 77(2): 508–519. doi:10.3758/s13414-014-0786-0.

Face features and face configurations both contribute to visual crowding

Hsin-Mei Sun and Benjamin Balas

Department of Psychology Center for Visual and Cognitive Neuroscience North Dakota State University

Abstract

Crowding refers to the inability to recognize an object in peripheral vision when other objects are presented nearby (Whitney & Levi, 2011). A popular explanation of crowding is that features of the target and flankers are combined inappropriately when they are located within an integration field and thus impair target recognition (Pelli, Palomares, & Majaj, 2004). However, it remains unclear which features of the target and flankers are combined inappropriately to cause crowding (Levi, 2008). For example, in a complex stimulus (e.g., a face), to what extent does crowding result from the integration of features at a part-based level or at the level of global processing of configural appearance? In this study, we used a face categorization task and different types of flankers to examine how much the magnitude of visual crowding depends upon similarity of face parts or similarity of global configurations. We created flankers with face-like features (e.g. the eyes, nose, and mouth) in typical and scrambled configurations to examine the impact of part appearance and global configuration on the visual crowding of faces. Additionally, we used “electrical socket” flankers that mimicked 1st-order face configuration but had only schematic features to examine the extent to which global face geometry impacted crowding. Our results indicate that both face parts and configurations contribute to visual crowding, suggesting that face similarity as realized under crowded conditions includes both aspects of facial appearance.

Keywords

Crowding; Face Perception; Configural Processing

Introduction

The presence of flanking stimuli substantially impairs the recognition of targets in the visual periphery. This phenomenon is known as crowding (Levi, 2008). Crowding has a number of distinct properties (Pelli, Palomares, & Majaj, 2004) that differentiate it from other superficially similar phenomena (e.g. masking). These include the scaling of the critical region for crowding to occur with eccentricity (Bouma, 1973), which is known as Bouma's Law, the impairment of recognition rather than detection (Pelli & Tillman, 2008), and the subjective appearance of the targets under conditions of crowding, which are typically reported as looking jumbled or mixed-up (Levi & Whitney, 2011; Pelli et al., 2004).

Crowding is also a very general phenomenon – similar properties hold for a wide range of visual qualities including color, shape, and motion (van der Berg, Roerdink, & Cornelissen, 2007). Crowding thus appears to be a key feature of human perception that possibly reflects a fundamental bottleneck limiting the fidelity with which complex scenes can be encoded (Levi, 2008) or alternatively, as a means of homogenizing appearance across distinct objects in a cluttered scene (Greenwood, Bex, & Dakin, 2010).

Crowding is often studied using relatively simple stimuli - Gabor patches for example (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001) - but crowding also occurs with complex objects (Wallace & Tjan, 2011). Using real-world objects, Wallace and Tjan (2011) demonstrated that “object crowding” had essentially the same properties observed using simpler stimuli, but that the crowding effect was weaker. Crowding strength (the negative impact of the flankers on target recognition) is sensitive to the similarity between targets and flankers (Kooi, Toet, Tripathy, & Levi, 1994) but clear similarity metrics are difficult to establish for complex targets, which makes it difficult to quantify how target-flanker similarity contributes to visual crowding when complicated objects are used. However, one can potentially use the extent to which crowding occurs given a particular pairing of targets and flankers as a means of characterizing the similarity between the target object and the flankers, as well as determining how the stimuli in the object array are encoded. For example, Martelli, Majaj and Pelli (2005) argued that the extent to which an object “self-crowds” – or contains discrete features that interfere with one another's appearance via crowding – can be used to determine whether or not holistic processing is applied to the object in question. Similarly, Louie, Bressler, and Whitney (2007) demonstrated that for upright target faces, inverted flankers had an appreciably lesser impact on target recognition than upright flankers. To rule out the possibility that their orientation effect was the result of changes in target/flanker similarity following inversion, they also investigated the impact of upright/inverted flankers on performance with an inverted target. This latter comparison did not yield a significant effect of flanker orientation, but nonetheless Louie et al. (2007) suggested that the effect of flanker orientation on target recognition could be the result of “holistic crowding,” or crowding that occurs after so-called holistic descriptors of face appearance are computed (see also Farzin, Rivera, & Whitney, 2009). Crowding, in these cases, is used as a tool to determine the vocabulary of recognition in some circumstances – if stimuli lead to visual crowding, this tells us that the visual system finds targets and flankers sufficiently alike to be confusable when stimulus appearance is combined. Similarly, critical features of complex stimuli can be manipulated to reveal what particular features within targets and flankers appear to be used by the visual system, as revealed by the extent to which the features in targets and flankers lead to perceptual confusion. Obviously it is more difficult to draw firm conclusions about the representations employed for object processing as object complexity increases (a problem that is not limited to visual crowding studies). The intersection of visual crowding and object recognition nonetheless provides important insights into the representations available to the observer for recognition when stimuli appear in dense, cluttered contexts.

In the current study, our goal was to investigate the visual crowding of faces to determine the contributions made to the crowding effect by both the discrete parts of a face (e.g., the eyes, nose, and mouth) and the global arrangement of parts into a face-like configuration

(e.g., the 1st-order configuration of a face). Both qualities of facial appearance are represented in the ventral object recognition pathway. The responses of early ERP and MEG components like the P100 (Hermann, Ehrlis, Ellgring, & Fallgatter, 2005) and the M100 (Liu, Harris, & Kanwisher, 2002) suggest that faces are represented primarily in terms of parts at early stages of visual processing. The appearance of parts, largely independent of their relative positions, modulates these responses, indicating that face configurations are not processed early in recognition. By comparison, later components of both ERP and MEG waveforms are more sensitive to the arrangement of parts within a face template. Both the N170 (Bentin, Allison, Puce, Perez, & McCarthy, 1996) and the M170 (Liu et al., 2002) exhibit profiles of sensitivity to face stimuli that suggest these later stages of processing are primarily dedicated to the encoding of global structure. Scrambled or isolated face parts typically do not yield as large a response at these components as the whole face template (though see Itier, Van Roon, and Alain (2011) for recent results describing the sensitivity of the N170 component to isolated eyes), and both responses remain relatively large when highly schematic or blurred faces are presented to observers (Flevaris, Robertson, & Bentin, 2008). These latter classes of stimuli have negligible local part structure, meaning that the responses observed must largely be driven by the global template of a face rather than a detailed representation of isolated features. These electrophysiological results are complemented by multiple fMRI and TMS studies of the fusiform face area (Pitcher, Walsh, Yovel, & Duchaine, 2007; Schlitz & Rossion, 2006) and the occipital face area (Pitcher, Walsh, & Duchaine, 2011), which appear to represent global face configuration and local part structure respectively.

Given that the discrete parts of a face and the global configuration of face stimuli are encoded at different stages of visual processing, we examined how each aspect of facial appearance contributed to the crowding of target faces by using different types of flanking stimuli. Recent theories of crowding generally agree that some sort of excessive integration of visual information including both the target object and the flankers leads to impaired recognition of the target – this presumptive integration has been described as a simple average of low-dimensional stimuli (Parkes et al., 2001), or as a representation of crowded arrays via texture processing (Balas, Nakano, & Rosenholtz, 2009) that can be applied to arbitrarily complex stimuli. In either case, a critical question is what visual information is being aggregated within the window of crowding. For complex objects like faces, the distinct representation of features and configurations at different stages of processing may mean that different properties of facial appearance impact visual crowding to varying degrees.

In a series of three experiments, we asked participants to categorize target faces in the visual periphery according to biological sex¹. The target faces were either presented in isolation, or flanked by face-like stimuli or Chinese characters, which we included as a non-face control

¹It is worth noting that studies have shown that certain facial features (e.g., the eyes, the eyebrows, the mouth, the face outline) play more influential roles in gender discrimination than do others (Brown, & Perrett, 1993; Dupuis-Roy, Fortin, Fiset, & Gosselin, 2009; Yamaguchi, Hirukawa, & Kanazawa, 1995). For example, Dupuis-Roy et al. (2009) had participants categorize the gender of a face presented behind a gray mask punctured by randomly located Gaussian apertures (the so-called “bubble mask”). They found that the availability of the eyes, the eyebrows, and the mouth was positively correlated with participants’ gender categorization performance, indicating an influential role of these facial features in gender categorization.

to measure crowding when a highly dissimilar object class was used. The face-like stimuli we used varied across our tasks to include line drawings of faces with recognizable facial features in a typical arrangement (Experiment 1), the same line drawings with the facial features scrambled such that the local appearance of parts was preserved but the global configuration was atypical (Experiment 2), and finally “electrical sockets” with simple shapes positioned in a face-like configuration such that local part appearance was not face-like, but face geometry was preserved (Experiment 3). To preview our results, we found that flankers with either face-like local part appearance (i.e., scrambled faces) or global configuration (i.e., electrical sockets) induced a crowding effect that was larger than that observed with Chinese characters, but smaller than that observed with intact line drawings of faces. While our design does not allow us to interpret our results directly in terms of specific neural stages, our goal was to reveal how these qualities of facial appearance contribute to the visual crowding of faces.

Experiment 1

In our first experiment, our goal was to establish that we could observe a significant crowding effect on face recognition using the line drawing stimuli that we wished to manipulate in subsequent experiments examining how part structure and face configurations contribute to crowding. In particular, we wished to compare the crowding effect induced with flankers comprised of intact line drawings of faces to that induced by flankers comprised of Chinese characters. These two categories of stimuli (face and non-face flankers) were intended to serve as reference points for the manipulated flankers used in subsequent tasks.

Methods

Participants—Twenty-five undergraduates (8 males) from North Dakota State University took part in this experiment for course credit. All participants were between the ages of 18-21 and reported either normal or corrected-to-normal vision. All participants were naïve to the purpose of the experiment. The experimental protocol was approved by the North Dakota State University Institutional Review Board for the protection of human participants in research, and written informed consent was obtained from all participants.

Apparatus—Participants were tested individually in a room with the light off. Stimuli were presented on an 18-inch CRT monitor with a refresh rate of 100 Hz and a display resolution of 1024×768 . Participants’ viewing distance to the monitor was stabilized at 60 cm with a HeadSpot chin rest. Responses were collected through the computer keyboard. Stimulus timing and response collection routines were implemented in the MATLAB Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Stimuli—Sixteen front-view Caucasian faces (8 male, 8 female) with neutral expressions were selected from the grayscale Face Recognition Technology (FERET) database and used as the target stimuli (see Appendix A). The flankers used in this task were comprised of Chinese characters and line drawings of faces produced with Adobe Illustrator (Figure 1). We created 6 exemplars of each flanker class and all of these stimuli were rendered entirely with black lines and matched to the dimensions of the target stimuli. From a viewing

distance of 60 cm, all stimuli subtended approximately 2.69° in height and 2.18° in width. The center-to-center spacing between target and flanker was about 2.52° in the flanked condition. During the experiment, all stimuli were presented on a medium-gray background and our stimuli were displayed in two-tone black and white.

Procedure and design—On each trial of the task, a fixation point appeared in the center of the screen for 500 ms, followed by a target face presented either at fixation (0 degrees eccentricity) or 6 degrees² to the left or right of fixation for 200 ms. Additionally, the target face was either presented in isolation or surrounded by flankers from either of the two classes (Chinese characters or face line drawings), which were either presented upright or inverted. Next, a blank screen was presented for 500 ms, followed by a question mark that remained visible on the screen until the participant responded (Figure 2). Participants were asked to categorize each target face as male or female by pressing the left or right arrow key on the computer keyboard. Participants were given unlimited time to respond; no feedback was provided after each response. The experiment used a 3 (target eccentricity: −6°, 0°, 6°) × 2 (flanker type: line-drawn face, Chinese character) × 2 (flanker orientation: upright, inverted) factorial design. Participants completed 32 trials per condition for a grand total of 384 trials with flankers present and an additional 192 trials (64 at each eccentricity) with no flankers. The experimental trials were preceded by 18 practice trials. Additionally, the experimental trials were presented randomly for each participant, in four blocks of 144 trials each.

Results

For each participant we calculated the proportion of correct responses in each condition (see Figure 3 and Appendix B). Percentages of correct responses were then submitted to a 2 × 2 × 2 repeated-measures ANOVA with target eccentricity (fovea or periphery, combining left and right offset), flanker type (Chinese characters or line-drawn face), and flanker orientation (upright or inverted) as within-subject factors. The analysis showed significant main effects of target eccentricity [$F(1,24) = 77.68, p < 0.001$] and flanker type [$F(1,24) = 44.92, p < 0.001$], but not flanker orientation [$F(1,24) = 0.26, p = 0.872$]. Importantly, there was a significant interaction between target eccentricity and flanker type [$F(1,24) = 20.41, p < 0.001$]. Paired-samples t-tests revealed that the interaction was driven by a significant difference between the different flanker conditions that was evident in the periphery [$t(24) = -7.92, p < 0.001$], but not in the fovea [$t(24) = -1.57, p = 0.130$]. That is, the crowding effect was stronger when a peripheral target face was surrounded by line drawings of faces compared to Chinese characters. There were no other significant main effects or interactions (all $ps > 0.17$).

Discussion

In Experiment 1 we established several important properties of our task that are necessary to examine the relative contribution of face parts and face configurations in our subsequent experiments. First, we demonstrated that our specific targets, flankers, and categorization

²We chose the 6° target eccentricity based on previous crowding studies. For example, Farzin et al. (2009) presented target faces at eccentricities of 0°, 3°, 6°, and 10°, and showed a significant crowding effect only when targets were presented at 6° eccentricity.

task were adequate to observe visual crowding, as evidenced by the effect of eccentricity in our data. The fact that poorer performance is observed when line-drawn face flankers surround the target in the periphery than when it appears in isolation eliminates the possibility that visual acuity is a limiting factor in our study. Second, we demonstrated that the similarity between the target and the flankers matters, in accord with previous results showing that crowding increases as target-flanker similarity gets higher (Bernard & Chung, 2011; Chung, Levi, & Legge, 2001; Kooi et al., 1994). Our participants had poorer categorization performance when a target face was flanked by line drawings of faces compared to Chinese characters. The Chinese characters we included in this task essentially do not lead to a measurable crowding effect³, suggesting that these flanking stimuli can be used as a reasonable lower-bound for target/flanker similarity and the subsequent effects on categorization performance under crowded conditions. As a result, we continue by comparing the impact of line-drawn face flankers to Chinese characters as we vary the parts and configurations of our flanking faces, and ultimately compare the impact of these manipulated flanking faces to one another. We do point out, however, that we did not observe the interaction between the orientation of face flankers and flanker type reported by Louie et al. (2007) and later in Farzin et al. (2009). Presently, we do not take the lack of replication in this task as any kind of referendum on these prior results, but at the very least it does suggest that the such flanker orientation effects are sensitive to stimulus and task parameters that varied between our study and previous reports. To further explore our main theme, we continue in Experiment 2 by examining the role that face parts play in crowding independent of the global configuration of face features.

Experiment 2

In our second experiment, we examined the extent to which the appearance of discrete face parts was sufficient to induce a crowding effect on target faces relative to non-face objects. By scrambling the arrangement of the eyes, nose, and mouth within the line drawings used as face flankers in Experiment 1, we preserved the structure of segmentable face features, but disrupted the 1st-order configuration of the face.

Methods

Participants—Twenty-five undergraduates (14 males) from North Dakota State University took part in this experiment for course credit. All participants were between the ages of 18-22 and reported either normal or corrected-to-normal vision. All participants were naïve to the purpose of the experiment and none had participated in Experiment 1. The North Dakota State University Institutional Review Board approved the study and all participants gave informed consent.

Apparatus, stimuli, procedure, and design—Experiment 2 was identical to Experiment 1, except that we created a new set of face-like flankers by rearranging the locations of the facial features of our original line-drawn faces (Figure 1). These scrambled

³Paired-samples t-tests comparing no flanker condition with Chinese character flanker condition in the periphery showed no significant differences ($[t(24) = 1.79, p = 0.09]$, $[t(24) = 0.34, p = 0.74]$, $[t(24) = 0.76, p = 0.46]$, for Experiment 1, Experiment 2, and Experiment 3, respectively).

faces thus had identical face parts (e.g., eyebrows, eyes, nose, mouth) as the line-drawn faces used in Experiment 1, but did not share their global configuration. Figure 4 shows the sequence of events in each trial and sample stimuli for different experimental conditions.

Results

We analyzed the proportion correct responses in each condition (see Figure 5 and Appendix B) using a $2 \times 2 \times 2$ repeated-measures ANOVA with the same within-subjects factors (target eccentricity, flanker type, flanker orientation) as described in Experiment 1. As in the previous experiment, the results revealed a main effect of target eccentricity [$F(1,24) = 37.18, p < 0.001$], a main effect of flanker type [$F(1,24) = 8.85, p < 0.01$], and an interaction between the two factors [$F(1,24) = 13.26, p = 0.001$]. Again, the main effect of eccentricity was driven by lower accuracy for targets presented in the periphery, and the main effect of flanker type was the result of lower accuracy for targets surrounded by scrambled faces. The interaction between target eccentricity and flanker type arose from a difference between scrambled face and Chinese character flankers that was only evident when targets were presented in the visual periphery [$t(24) = -4.19, p < 0.001$], but not in the fovea [$t(24) = 0.27, p = 0.789$]. As in Experiment 1, participants had worse target categorization performance when a peripheral target face was surrounded by scrambled faces compared to Chinese characters. No other main effects or interactions were significant (all $ps > 0.284$).

Discussion

The data from Experiment 2 indicate that flankers with scrambled face features induce a crowding effect on face recognition that is larger than the effect induced by Chinese characters. This suggests that breaking the 1st-order configuration of face parts does not sufficiently impact the similarity between target faces and scrambled line-drawn faces to lessen the effect of crowding to levels that are comparable to those achieved with highly dissimilar flankers. One way to interpret these results (that is admittedly speculative) is that crowding may occur at a level of representation where face parts are processed more or less in isolation from their arrangement into a global gestalt, like the occipital face area (Liu, Harris, & Kanwisher, 2010; Nichols, Betts, & Wilson, 2010; Pitcher et al., 2011). Obviously we cannot unequivocally conclude that our results have such a clear neural interpretation, but we raise this interpretation as an interesting possibility for further consideration given that the architecture of face processing in the ventral visual system has been elaborated via neuroimaging studies. Another account that these data are consistent with is the representation of target and flanker appearance in terms of summary statistics (Balas et al., 2009). Balas et al.'s model of visual crowding assumes that the entire stimulus array presented to observers in the periphery is summarized by the visual system via a texture-like code for appearance. This code is lossy, but contains sufficient information to constrain the class of potential targets (and flankers) enough for a range of categorization tasks to be accomplished (Rosenholtz, 2011). In particular, a great deal of spatial information is lost when representing crowded arrays via texture statistics. In terms of the current data, our scrambled faces are largely commensurate with the equivalence class of stimuli that a texture code might impose on the appearance of targets and flankers in this task. Our result thus may also reflect the fact that at early stages of visual perception the description of the targets and flankers in a crowded stimulus array is largely blind to the global arrangement of

parts since that information has probably been lost in the encoding of image appearance via a statistical code (see Figure 10 in Balas et al. (2009) for an example of how texture representations can fail to constrain the position of more complex features within an image). Again, we offer this as one possible way to interpret these results in terms of mechanisms of visual processing that are instantiated in the ventral visual stream, even though our design does not permit us to draw firm conclusions about the neural processes that underlie our behavioral results. The key inference that the data from Experiment 2 allows us to make is that disrupting global configuration within a face does not sufficiently disrupt the appearance of flankers to reduce crowding to the level achieved with highly non-face-like flankers. While this does not allow us to infer much about specific neural mechanisms, it does tell us about the similarity relationships that impact crowding when face stimuli are targets, which is an important contribution in its own right. To complement these results, we therefore continue in Experiment 3 by determining the extent to which global face configuration is sufficient to induce a crowding effect that is larger in magnitude than that achieved using flankers from an unrelated stimulus class.

Experiment 3

In Experiment 3, we examined the magnitude of visual crowding when target faces were flanked by Chinese characters or schematic “electrical socket” faces. This latter stimulus class had discrete parts comprised of basic shapes (e.g. solid rectangles) that were arranged into a crude face template such that surrogate eyes and nose features were located in an approximation of a human face. These stimuli were designed to complement the scrambled face flankers used in Experiment 2 insofar as they offered a means of assessing how similarity at the level of global configuration contributed to the crowding of face targets when flanker face parts bore essentially no resemblance to those in the target faces.

Methods

Participants—Twenty-five undergraduates (14 males) from North Dakota State University took part in this experiment for course credit. All participants were between the ages of 18-22 and reported either normal or corrected-to-normal vision. All participants were naïve to the purpose of the experiment and none had participated in Experiment 1 or Experiment 2. The North Dakota State University Institutional Review Board approved the study and all participants gave informed consent.

Apparatus, stimuli, procedure, and design—Experiment 3 was identical to Experiment 1, except that we created a new set of face-like flankers by drawing schematic “electrical socket” faces in Adobe Photoshop (Figure 1). We generated 6 flankers from this class according to the arrangement of several actual electrical outlets from various countries. Face parts, such as they are, were drawn using simple geometrical shapes including oriented rectangles and circles. Before starting the experiment, we explicitly told participants that the task required them to categorize target faces flanked by either Chinese characters or electrical sockets from different countries. Figure 6 shows the trial sequence and sample stimuli.

Results

As in Experiments 1 and 2, we analyzed the proportion correct responses (see Figure 7 and Appendix B) in each condition using a $2 \times 2 \times 2$ repeated-measures ANOVA with the same within-subjects factors (target eccentricity, flanker type, flanker orientation). As in both previous experiments, this analysis revealed a main effect of target eccentricity [$F(1,24)=48.79, p<0.001$], and a main effect of Flanker type [$F(1,24) = 24.47, p < 0.001$], as well as an interaction between these two factors [$F(1,24) = 23.82, p < 0.001$]. Both main effects were the result of target eccentricity and flanker type affecting target categorization in the same manner as observed in Experiment 1: Targets in the periphery were harder to categorize, as were targets flanked by face-like stimuli. Also, the interaction between these two factors was the result of a difference between the Chinese character and the electrical socket flankers that was only evident in the periphery [$t(24) = -6.26, p < 0.001$], but not in the fovea [$t(24) = -0.19, p = 0.852$]. No other main effects or interactions reached significance (all $ps > 0.362$).

Given our stated goal of assessing the relative contributions of face parts and face configurations to the visual crowding of face targets, we also carried out a combined analysis of the data from all three face-like flanker conditions to directly compare the magnitude of the crowding effect across these conditions. We combined the data from the trials where face-like flankers were used in Experiments 1, 2, & 3 into a mixed-design 2×3 ANOVA with target eccentricity (fovea, periphery) as the within-subjects factor and flanker type (original line drawings, scrambled line drawings, electrical sockets) as a between-subjects factor. The results showed a significant main effect of target eccentricity [$F(1,72) = 179.24, p < 0.001$], but no significant main effect of flanker type [$F(2,72) = 1.139, p = 0.33$]. However, the interaction between target eccentricity and flanker type was significant [$F(2,72) = 4.78, p = 0.011$]. Planned comparisons revealed that in the periphery, crowding was stronger when a target face was flanked by intact line drawings of faces (Experiment 1) compared to scrambled faces (Experiment 2) that shared the same face parts [$t(48) = -2.30, p = 0.013$, one-tailed independent samples t-test⁴]. The difference between the intact line drawings of faces and the electrical socket flankers in the periphery was also significant [$t(48) = -1.77, p = 0.042$, one-tailed independent samples t-test]; participants showed more crowding when a target face was surrounded by intact line drawings of faces compared to electrical sockets (Experiment 3) that shared their global configuration. Finally, the difference between the scrambled face flankers and the electrical sockets was not significant ($p = 0.466$).

General Discussion

The results of our three experiments support the conclusion that face categorization performance under crowded conditions is sensitive to the global arrangement of parts within a face and the appearance of discrete face parts themselves, but that neither aspect of facial appearance was so critical to target-flanker similarity that selective disruption of these properties would reduce crowding to the levels observed when non-face objects were used.

⁴We used one-tailed p -values for this and the next test because we hypothesized that line-drawn face flankers, which retained both global and local facial features, would cause more crowding compared to either scrambled face or electrical socket flankers.

Both discrete face parts in a non-canonical arrangement (Experiment 2) and schematic features arranged in a typical face configuration (Experiment 3) induced crowding effects that were larger than those resulting from Chinese character flankers. One speculative conjecture regarding our results is that visual crowding may be realized both at the level of the representation of face parts and at the level of holistic representations of facial appearance. In direct comparison to one another, performance with intact face flankers compared to our “electrical socket” stimuli did differ significantly, suggesting that local part appearance plays an important role in determining the magnitude of crowding for face stimuli.

To what extent can we interpret these results solely in terms of low-level image similarity? The difference between intact face flankers and “electrical sockets” is largely consistent with a general account of crowding as the by-product of texture-like descriptors (Balas et al., 2009) that summarize appearance via feature histograms that capture marginal and joint statistics within some spatial neighborhood. Such a representation would likely be sufficient to encode the difference in appearance between the targets and our “electrical sockets” since these differ fairly broadly in terms of orientation features and correlations between local edges. By contrast, the difference between performance in the intact face flanker condition and the scrambled face flanker condition, makes it harder to accept a pure summary-statistic account of the data, since such models are highly unlikely to discriminate between scrambled and intact flankers – the scale at which feature correlations are measured makes it unlikely that the texture code used by Balas et al. (2009) would reliably constrain appearance such that these sets of flankers would differ. We thus suggest that this particular definition of similarity which relies on low-level features (edges and their correlations) rather than on intermediate or high-level features (eyes, nose, mouth, or a global template of the whole face) does not account for all the data. Presently, we therefore interpret our data as evidence that face similarity as computed under crowded conditions includes contributions of both the global layout of features within a face (i.e., 1st-order face configuration) as well as the appearance of local features (e.g., the eyes, nose, and mouth) such that target-flanker similarity is not dominated by either characteristic of face appearance, but is instead sensitive to both. Crowding thus may impact stages of processing following the computation of features that constrain these properties of the face.

One interesting issue to consider is why we did not obtain the interaction between orientation and flanker type that one would expect from prior studies examining the impact of flanker face orientation on crowding (Farzin et al., 2009; Louie et al., 2007). To the extent that our data might be interpreted as evidence that crowding occurs after local feature appearance and global configuration is computed, we might expect that flanker inversion should impact performance due to disruptions in how global configuration is computed in inverted faces. The fact that we do not observe such effects in our task may be due to a range of factors - the larger dissimilarity between target faces and flankers in our experiments, or stimulus differences between our target faces and those used in previous reports. As such, we interpret the absence of the upright/inverted flanker effect in our data as little more than a consequence of the particular stimuli we used to manipulate local and global features in this task, or perhaps a consequence of task design (number of trials, subjects, etc.).

How do our results relate to previous investigations of how high-level descriptors of stimulus appearance impact crowding? We suggest that studies comparing crowding to ensemble processing (or “seeing sets” of objects (Ariely, 2001)) are an interesting point of comparison for our results, since these studies are also concerned with determining what information has been computed at the stage where crowding affects performance. In particular, computation of the average stimulus appearance may or may not be impacted by crowding, depending on the types of stimuli. For example, Banno and Saiki (2012) revealed a deleterious impact of nearby flankers on calculation of average circle size, suggesting that statistical information about the entire array cannot pass through or circumvent the crowding bottleneck intact. However, studies have also shown that average emotional expression can be extracted from crowded arrays of faces (Fischer & Whitney, 2011; Haberman & Whitney, 2007; 2009; Kouider, Berthet, & Faivre, 2011), suggesting that the computation of the average may belie the impact of crowding on earlier, less holistic stages of processing. Comparing crowding performance in face recognition tasks to results obtained with simpler stimuli (e.g. circles), it is therefore important to consider that the visual crowding of faces may have distinct properties that result from differential processing within an extended network of face processing loci (Haxby, Hoffman, & Gobbini, 2000). That is, at some stages of the ventral visual pathway, a target might not be harder to recognize in the presence of some specific flankers because the encoding of the target and the flankers at that stage may be sufficiently different so that the appearance of the target may be sufficiently constrained for accurate recognition. At another stage, this may not be the case. The representation of face parts and face configurations at distinct stages of neural processing thus may make it possible for crowding to be realized differently at distinct stages of visual processing. Again, we emphasize that our current design does not allow us to draw conclusions about the neural implementation of crowding at these putative stages of face processing, but we offer this possibility here as an interesting conjecture.

Presently, by dissociating featural and configural information we have made it possible to observe the extent to which global information is preserved and contributes to perception during crowding (Fischer & Whitney, 2011) while also finding evidence that individual features can impact the crowded percept despite disruption of global flanker appearance (Faivre & Kouider, 2011). The crowded perception of complex objects (faces, in particular) thus appears to be determined by integration of both the local and global features of an object.

Acknowledgements

BB was supported by COBRE Grant P20 GM103505 from the National institute for General Medical Studies (NIGMS) and NSF EPSCoR Grant #EPS-0814442. The authors would like to thank four anonymous reviewers for their helpful comments on the earlier versions of this manuscript. The authors would also like to thank Christopher Tonsager for his assistance with data collection.

Appendix

Target faces used in Experiments 1-3.

Male faces

Female faces

Appendix A.

Appendix

Appendix B

Mean percentage of accuracy for each experimental condition in Experiments 1-3. The standard errors of the means are in parentheses.

Experiment 1

Target Eccentricity	No Flankers	CC Flankers		Face Flankers	
		Upright	Inverted	Upright	Inverted
Fovea	92.9 (1.0)	95.3 (0.8)	95.4 (0.9)	94.5 (1.3)	93.2 (1.5)
Periphery	82.8 (1.6)	80.6 (2.1)	81.9 (1.9)	74.1 (2.3)	73.3 (2.2)

Experiment 2

Target Eccentricity	No Flankers	CC Flankers		SF Flankers	
		Upright	Inverted	Upright	Inverted
Fovea	92.1 (1.3)	91.5 (1.3)	92.8 (1.1)	92.7 (1.5)	92.0 (1.2)
Periphery	84.8 (1.7)	84.8 (1.8)	84.2 (1.9)	80.5 (2.2)	81.0 (2.1)

Experiment 3

Target Eccentricity	No Flankers	CC Flankers		ES Flankers	
		Upright	Inverted	Upright	Inverted
Fovea	92.9 (0.8)	93.7 (1.2)	92.7 (1.2)	92.7 (1.3)	93.4 (0.8)
Periphery	84.5 (1.4)	83.8 (1.3)	83.8 (1.8)	78.8 (1.9)	78.3 (1.8)

Note. CC = Chinese Character

Note. CC = Chinese Character; SF = Scrambled Face

Note. CC = Chinese Character; ES = Electrical Socket

References

- Ariely D. Seeing sets: Representation by statistical properties. *Psychological Science*. 2001; 12:157–162. [PubMed: 11340926]
- Balas B, Nakano L, Rosenholtz R. A summary-statistical model of peripheral vision explains visual crowding. *Journal of Vision*. 2009; 9(12)(13):1–9.
- Banno H, Saiki J. Calculation of the mean circle size does not circumvent the bottleneck of crowding. *Journal of Vision*. 2012; 12(11):13, 1–15. [PubMed: 23090611]
- Bentin S, Allison T, Puce A, Perez E, McCarthy G. Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*. 1996; 8:551–565. [PubMed: 20740065]
- Bernard J-B, Chung STL. The dependence of crowding on flanker complexity and target-flanker similarity. *Journal of Vision*. 2011; 11(8):1, 1–16. [PubMed: 21730225]
- Bouma H. Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*. 1973; 13:767–782. [PubMed: 4706350]
- Brainard D. The Psychophysics Toolbox. *Spatial Vision*. 1997; 10:433–436. [PubMed: 9176952]
- Brown E, Perrett DI. What gives a face its gender? *Perception*. 1993; 22:829–840. [PubMed: 8115240]
- Chakravarthi R, Cavanagh P. Recovery of a crowded object by masking the flankers: Determining the locus of feature integration. *Journal of Vision*. 2009; 9(10):1–9.
- Chung S, Levi D, Legge G. Spatial frequency and contrast properties of crowding. *Vision Research*. 2001; 41:1833–1850. [PubMed: 11369047]
- Dupuis-Roy N, Fortin I, Fiset D, Gosselin F. Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*. 2009; 9(2):1–8. [PubMed: 19271920]
- Faivre N, Kouider S. Multi-feature objects elicit nonconscious priming despite crowding. *Journal of Vision*. 2011; 11(3):1–10.
- Farzin F, Rivera SM, Whitney D. Holistic crowding of Mooney faces. *Journal of Vision*. 2009; 9(6):18, 11–15. [PubMed: 19761309]
- Fischer J, Whitney D. Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology*. 2011; 106:1389–1398. [PubMed: 21676930]
- Flevaris AV, Robertson LC, Bentin S. Using spatial frequency scales for processing face features and face configuration: An ERP analysis. *Brain Research*. 2008; 1194:100–109. [PubMed: 18190897]
- Greenwood JA, Bex PJ, Dakin SC. Crowding changes appearance. *Current Biology*. 2010; 20:496–501. [PubMed: 20206527]

17. Haberman J, Whitney D. Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*. 2007; 17:R751–R753. [PubMed: 17803921]
18. Haberman J, Whitney D. Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*. 2009; 35:718–734. [PubMed: 19485687]
19. Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends in Cognitive Science*. 2000; 4:223–233.
20. Hermann MJ, Ehrls AC, Ellgring H, Fallgatter AJ. Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission*. 2005; 112:1073–81. [PubMed: 15583954]
21. Itier R, Van Roon P, Alain C. Species sensitivity of early face and eye processing. *Neuroimage*. 2011; 54:705–713. [PubMed: 20650321]
22. Kooi FL, Toet A, Tripathy SP, Levi DM. The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*. 1994; 8(2):255–279. [PubMed: 7993878]
23. Kouider S, Berthet V, Faivre N. Preference is biased by crowded facial expressions. *Psychological Science*. 2011; 22(2):184–189. [PubMed: 21233479]
24. Levi DM. Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*. 2008; 48:635–654. [PubMed: 18226828]
25. Liu J, Harris A, Kanwisher N. Stages of processing in face perception: An MEG study. *Nature Neuroscience*. 2002; 5(9):910–6.
26. Liu J, Harris A, Kanwisher N. Perception of face parts and face configurations: an fMRI study. *Journal of Cognitive Neuroscience*. 2010; 22:203–211. [PubMed: 19302006]
27. Louie EG, Bressler DW, Whitney D. Holistic crowding: selective interference between configural representations of faces in crowded scenes. *Journal of Vision*. 2007; 7(2):1–11. [PubMed: 18217839]
28. Martelli M, Majaj N, Pelli D. Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*. 2005; 5(1):6, 58–70.
29. Maurer D, Le Grand R, Mondloch CJ. The many faces of configural processing. *Trends in Cognitive Sciences*. 2002; 6:255–260. [PubMed: 12039607]
30. McKone E. Isolating the special component of face recognition: Peripheral identification and a Mooney face. *Journal of Experimental Psychology, Learning, Memory, and Cognition*. 2004; 30:181–197.
31. Nichols DF, Betts LR, Wilson HR. Decoding of faces and face components in face-sensitive human visual cortex. *Frontiers in Psychology*. 2010; 1:28. [PubMed: 21833198]
32. 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.
33. Pelli D. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*. 1997; 10:437–442. [PubMed: 9176953]
34. Pelli DG, Tillman KA. The uncrowded window of object recognition. *Nat Neurosci*. 2008; 11(10):1129–1135. [PubMed: 18828191]
35. Pelli DG, Palomares M, Majaj NJ. Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*. 2004; 4:1136–1169. [PubMed: 15669917]
36. Pitcher D, Walsh V, Yovel G, Duchaine B. TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology*. 2007; 17(18):1568–1573. [PubMed: 17764942]
37. Pitcher D, Walsh V, Duchaine B. The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*. 2011; 209:481–493. [PubMed: 21318346]
38. Rosenholtz, R. What your visual systems sees where you are not looking.. In: Rogowitz, BE.; Pappas, TN., editors. *Proceedings of SPIE: Human vision and Electronic Imaging XVI*. San Francisco: 2011.
39. Schiltz C, Rossion B. Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage*. 2006; 32:1385–1394. [PubMed: 16870475]

40. van den Berg R, Roerdink JB, Cornelissen FW. On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision*. 2007; 7(2):14, 1–11. [PubMed: 18217829]
41. Wallace JW, Tjan BS. Object Crowding. *Journal of Vision*. 2011; 11(6):1–17.
42. Whitney D, Levi DM. Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*. 2011; 15:160–168. [PubMed: 21420894]
43. Yamaguchi MK, Hirukawa T, Kanazawa S. Judgment of gender through facial parts. *Perception*. 1995; 24:563–575. [PubMed: 7567430]

Chinese Characters
(Exps. 1-3) Line-drawn Faces
(Exp. 1) Scrambled Faces
(Exp. 2) Electrical Sockets
(Exp. 3) **Figure 1.**

Example stimulus arrays from Experiments 1, 2, and 3. During the experiment, flanking stimuli were presented either upright or inverted.

Figure 2.

Left: Illustration of a trial used in Experiment 1. Right: Example stimuli for different combinations of flanker type and orientation. All images depict trials in which the target face and its flankers were presented to the left of fixation.

Figure 3.

Average proportion correct across participants in different experimental conditions of Experiment 1. Error bars represent standard errors of the means. (CC: Chinese character)

Figure 4.
The trial sequence and sample stimuli in Experiment 2.

Figure 5.

Average proportion correct across all participants in different experimental conditions of Experiment 2. Error bars represent standard errors of the means. (CC: Chinese character; SF: scrambled face)

Figure 6.
The trial sequence and sample stimuli in Experiment 3.

Figure 7.

Average proportion correct across participants in different experimental conditions of Experiment 3. Error bars represent standard errors of the means. (CC: Chinese character; ES: electrical socket)