Holistic crowding: Selective interference between configural representations of faces in crowded scenes

Elizabeth G. Louie
The Department of Psychology, University of California, Davis, CA, USA

David W. Bressler
The Center for Mind and Brain, University of California, Davis, CA, USA

David Whitney
The Department of Psychology, University of California, Davis, CA, USA, & The Center for Mind and Brain, University of California, Davis, CA, USA

It is difficult to recognize an object that falls in the peripheral visual field; it is even more difficult when there are other objects surrounding it. This effect, known as crowding, could be due to interactions between the low-level parts or features of the surrounding objects. Here, we investigated whether crowding can also occur selectively between higher level object representations. Many studies have demonstrated that upright faces, unlike most other objects, are coded holistically. Therefore, in addition to featural crowding within a face (M. Martelli, N. J. Majaj, & D. G. Pelli, 2005), we might expect an additional crowding effect between upright faces due to interference between the higher level holistic representations of these faces. In a series of experiments, we tested this by presenting an upright target face in a crowd of additional upright or inverted faces. We found that recognition was more strongly impaired when the target face was surrounded by upright compared to inverted flanker (distractor) faces; this pattern of results was absent when inverted faces and non-face objects were used as targets. This selective crowding of upright faces by other upright faces only occurred when the target–flanker separation was less than half the eccentricity of the target face, consistent with traditional crowding effects (H. Bouma, 1970; D. G. Pelli, M. Palomares, & N. J. Majaj, 2004). Likewise, the selective interference between upright faces did not occur at the fovea and was not a function of the target–flanker similarity, suggesting that crowding-specific processes were responsible. The results demonstrate that crowding can occur selectively between high-level representations of faces and may therefore occur at multiple stages in the visual system.

Keywords: vision, perception, awareness, face recognition, ensemble, spatial, lateral, masking, object

Citation: 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://journalofvision.org/7/2/24/, doi:10.1167/7.2.24.

Introduction

We have all experienced the difficulty of trying to locate a face in a crowd. The fact that a face is much easier to recognize when it is located in the central visual field than when it is located in the periphery is not entirely due to poor visual acuity in the periphery, but also due to the presence of internal as well as surrounding features that interfere with the identification of the target face. This effect is called crowding (Bouma, 1970; Field, Hayes, & Hess, 1993; He, Cavanagh, & Intriligator, 1996; Intriligator & Cavanagh, 2001; Latham & Whitaker, 1996; Levi, Klein, & Aitsebaomo, 1985; Martelli, Majaj, & Pelli, 2005; Pelli, Palomares, & Majaj, 2004; Strasburger, Harvey, & Rentschler, 1991; Toet & Levi, 1992; Westheimer & Hauske, 1975). Unlike traditional masking, in which a signal (e.g., the face) is rendered invisible, crowding occurs when the signal is still visible but its features blend with its neighbors (Martelli et al., 2005; Pelli et al., 2004). Integrating surrounding features with those of the signal results in the inability to scrutinize or identify the target. According to most models, crowding occurs because of interference or pooling among low-level features, which likely happens at a single, relatively early stage in visual processing (Ariely, 2001; Chung, Levi, & Legge, 2001; He, Cavanagh, & Intriligator, 1997; Levi et al., 1985; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pelli et al., 2004). To date, however, it has not been tested whether crowding can occur selectively at higher levels in the visual system—only crowding of low-level features has been demonstrated. The possibility therefore remains that crowding may operate at multiple levels in the visual

doi: 10.1167/7.2.24 Received October 20, 2006; published November 26, 2007 ISSN 1534-7362 © ARVO

system; for example, even between high-level representations of objects.

In this study, we tested whether crowding can occur between high-level, holistic representations of objects, not just between low-level features as past research has found. To address this question, we used upright and inverted faces as stimuli. It is well established that recognition of an upright face is not necessarily based on the processing of its individual features (featural processing). Rather, we tend to identify upright faces holistically, or by analyzing the configuration or relations between these features (Boutet & Chaudhuri, 2001; Farah, Tanaka, & Drain, 1995; Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Grand, & Mondloch, 2002; Tanaka & Farah, 1993; Thompson, 1980; Yin, 1969; Young, Hellawell, & Hay, 1987). Moreover, McKone and colleagues (McKone, Martini, & Nakayama, 2001; Robbins & McKone, 2003, 2007) have convincingly shown that holistic processing cannot be learned for inverted faces or non-face objects, which provides further evidence that the processing of upright faces is distinct from that of inverted faces (though cf. Carey, 1992; Diamond & Carey, 1986). Thus, the presence of an inversion effect is a reliable indicator of holistic processing of upright faces.

If crowding occurs selectively between the configural representations of upright faces, and not just between the features of the faces themselves (Martelli et al., 2005), we would expect greater impairment of face recognition when an upright target face is surrounded by similarly configured upright faces compared to when it is surrounded by inverted faces. On the other hand, we would not expect the same pattern of results to hold true for non-face target objects (e.g., houses) or for inverted target faces. To test these predictions, we measured face and house recognition in a crowded display of other faces and houses.

Methods

Four experienced psychophysical subjects with normal or corrected-to-normal visual acuity participated in the experiments. All experiments were approved by the human subjects review board at UC Davis. Stimuli were presented on a gamma-corrected, linearized, high-resolution CRT monitor (Sony Multiscan G520, 1024 × 768 pixels, 120 Hz refresh) using an Apple G4 Power Macintosh with OS9 running Vision Shell (www.visionshell.com). Participants were seated in a dark soundproof room with a chin rest placed 49 cm from the screen.

Stimuli

Stimuli were 30 houses and 30 male faces with neutral expressions, drawn from Dr. Ken Nakayama's Harvard Face Database, with permission and consent. Using Adobe Photoshop 8.0, all stimuli were grayscale filtered, noise filtered (10%), and band-pass spatial frequency filtered with cutoffs at 1 cycle/pixel and 0.2 cycles/pixel. All stimuli were edited so that the main features fit inside an oval window of 3.51 deg visual angle wide and 4.82 deg high; the outlines of the stimuli (the edges of the faces and houses) were not visible.

Experiment 1: Upright faces among upright and inverted face flankers

Three conditions were presented. In the upright flankers condition (Figure 1a), a central face, which could or could not be the target face, was surrounded by six upright flanker faces, creating an array of faces; an array was presented on both sides of the fixation bull's-eye (11.58 deg center-to-center distance). In the inverted flankers condition, the array was composed of an upright central face surrounded by six inverted flanker faces (Figure 1b). In the third condition, the central face was presented without the surrounding flankers (Figure 1c). The array of crowding faces (the flankers) were presented on an imaginary oval surrounding each central face (5.24 deg vertical distance between the midpoint of the central face and the oval, and 3.98 deg horizontal distance between the midpoint of the central face and the oval). The spacing between each flanker was fixed, but the absolute position of the flankers on the oval was randomized on each trial. The average feature-to-feature (e.g., nose-to-nose) distance between the central and surrounding faces was equal for upright and inverted flankers. In addition to the noise added by Photoshop, one of five levels of random dot noise was added to the central face in both arrays on every trial by increasing or decreasing the brightness of each pixel by a random amount within ±31 cd/m^2, which ensured that recognition performance did not reach ceiling or floor (Figure 1d).

On each trial, the stimuli were presented for 400 ms. A central target face was presented on 50% of the trials in the array on either the left or right side of the screen. While continuously fixating on the bull's-eye, subjects performed a 3 alternative forced-choice (3 AFC) task to indicate whether the target face was on the left or right side of the screen, or was not present. This is comparable to a simultaneous detection and identification experiment in which the stimulus was not present in one condition (Green, Weber, & Duncan, 1977; Starr, Metz, Lusted, & Goodenough, 1975). Prior to the main experiment, each subject participated in a minimum of 2400 practice trials to reach asymptotic recognition levels and become familiarized with the target face (in 150 trial blocks; the first four blocks included feedback for missed responses). During these practice runs, an adaptive procedure was used to manipulate the level of noise randomly superimposed on each central face such that in the lowest noise

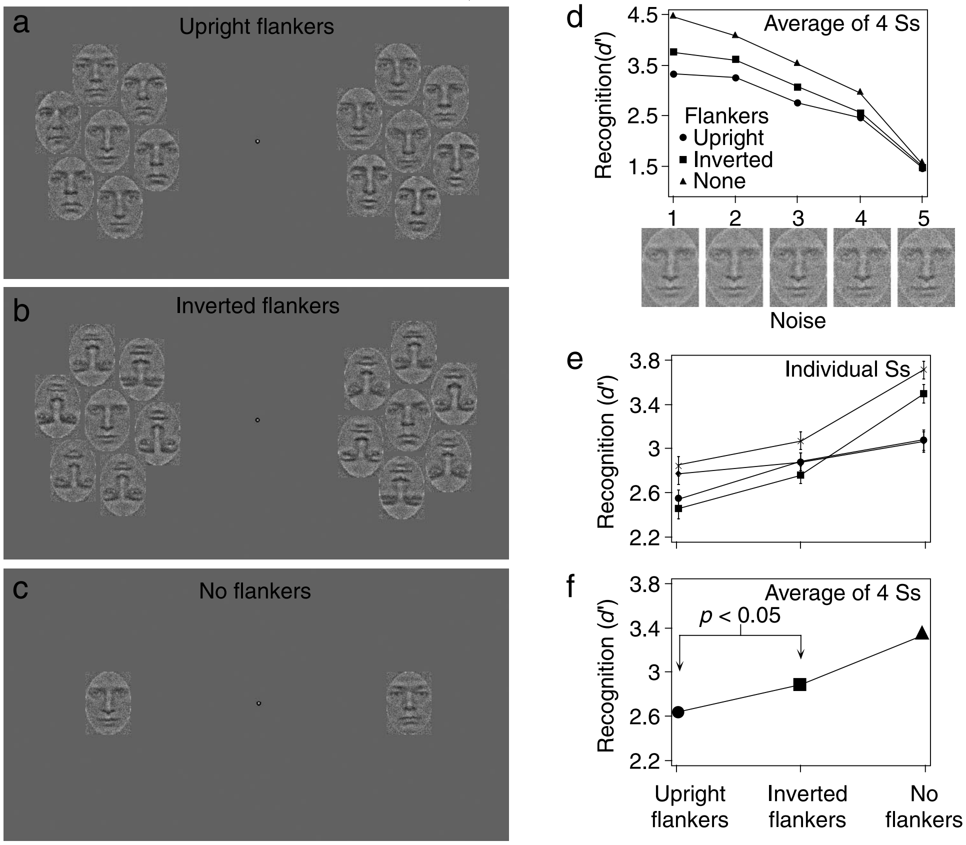

Figure 1. Three crowding conditions were presented in Experiment 1: upright flankers (a), inverted flankers (b), or no flankers (c); the central face could or could not be the target face. Subjects indicated whether the target face (Figure 1a, left side) was on the left or right side of the fixation point or not present. Across all conditions, recognition was impaired with increasing noise added to the stimuli (d). Despite the eccentric location of the target face (c), all subjects were able to recognize and discriminate it, consistent with previous studies (McKone, 2004). Panel e shows mean discrimination for four individual subjects (symbols represent separate subjects). Averaged across six subjects (the four original subjects plus two new subjects from Experiment 4), there was a significant reduction in *d*-prime in the upright versus inverted flanker conditions ($t(5) = 3.6$, $p < 0.01$, $r = 0.95$; Wilcoxon test, $Z = -1.99$, $P = 0.023$). This demonstrates that recognition was significantly and selectively impaired in the upright flanker condition (f). Error bars, within subjects ±*SEM*.

condition accuracy was above 90%, while in the highest noise condition the accuracy was between 50% and 60% (chance performance was 33%). In the main experiment, subjects participated in 10 separate sessions, with 300 trials per session (3 flanker conditions × 5 noise levels × 20 trials per condition), for a total of 3000 trials per subject. No feedback was provided. For each of the 10 sessions, *d*-prime was calculated for each condition as an indicator of face recognition sensitivity, as in multiple response classification methods (Haase & Fisk, 2001; MacMillan & Creelman, 2004); statistics were calculated on the mean *d*-primes.

Experiment 2: Upright houses among upright and inverted house flankers

The methods in this experiment were identical to those in Experiment 1, except the stimuli were images of houses (Figure 2a). Each subject from Experiment 1 completed a total of 2100 practice trials and 10 sessions (3000 trials) of Experiment 2.

Experiment 3: Inverted faces among upright and inverted face flankers

The methods in this experiment were identical to those in Experiment 1, except the central faces were inverted (Figure 3a). Three of the four subjects from Experiment 1 completed a total of 2400 practice trials and 10 sessions (3000 trials) of Experiment 3.

Experiment 4: Upright face crowding as a function of eccentricity

Several studies have demonstrated that crowding follows a unique half-eccentricity rule: Crowding occurs as

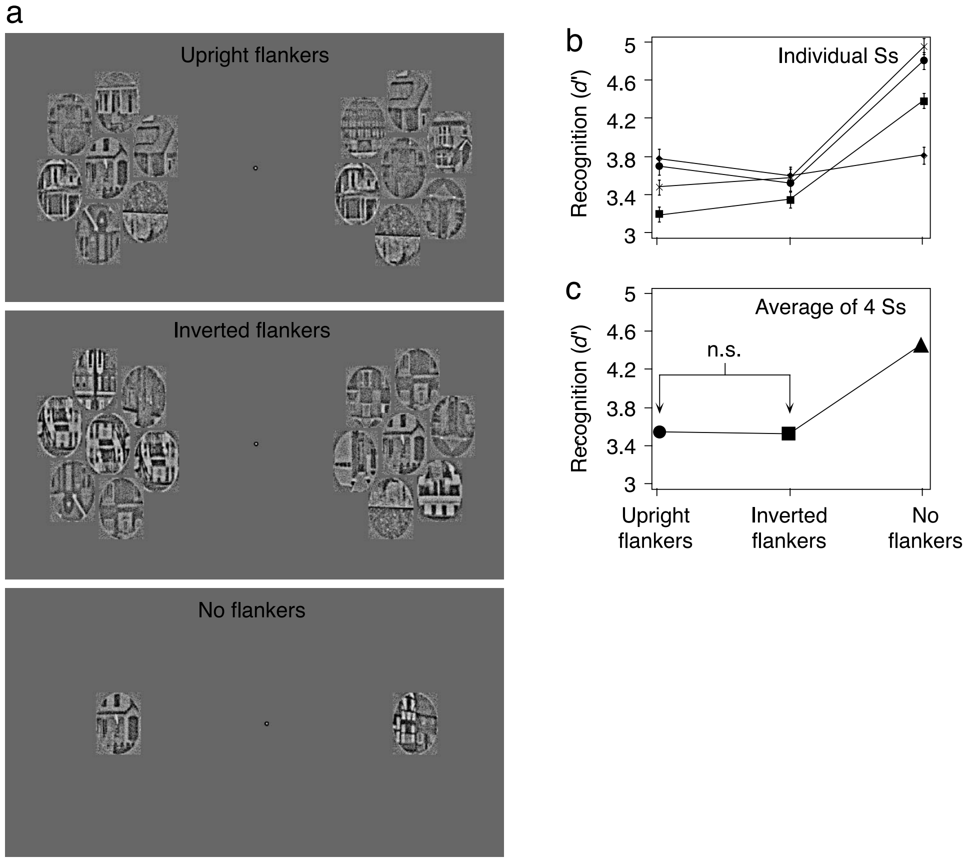

Figure 2. The conditions and tasks in Experiment 2 remained the same as those in Experiment 1, except houses were used as stimuli (a); the target house is located on the left in the top panel. Without crowding, each of the four subjects was easily able to identify the target house (b; each symbol type represents a separate subject). Panel c shows the average recognition across the four subjects. Although there was a difference between the no flanker and inverted flanker conditions (c), indicating that crowding was effective, there was no significant or selective difference between the upright and inverted flanker conditions ($F(1, 3) = 0.075$, $p > 0.05$, $\text{eta}^2 = 0.02$). Error bars, within subjects ±*SEM*.

long as the target–flanker separation is less than half the eccentricity of the target (Bouma, 1970; Martelli et al., 2005; Pelli et al., 2004). To ensure that the effect we observed is due to crowding and not masking, we repeated Experiment 1 at two additional eccentricities. One subject from Experiment 1 and two naïve subjects each participated. In one condition, targets were presented at 6.13 deg eccentricity, and subjects participated in 10 sessions (3000 trials); all other experiment details were identical to Experiment 1. In the second condition, targets were presented foveally (0 deg eccentricity). All stimulus details were identical to Experiment 1, but the task was a temporal version of the spatial simultaneous detection and identification task in Experiment 1: An array was flashed for 400 ms, followed by a 400-ms ISI, and a second array was flashed for 400 ms (Green et al., 1977; Starr et al., 1975). Subjects performed a 3 AFC task in which they indicated whether the target face appeared in the first interval, the second interval, or did not appear at all.

Experiment 5: Famous face recognition at eccentric locations

To determine whether observers can recognize an isolated face at an eccentric location, we tested the ability of seven naïve subjects to recognize upright and inverted famous faces in isolation at the three eccentricities presented in Experiments 1 and 4 (0 deg, 6.13 deg, and 11.58 deg from fixation). Images of 50 celebrities (25 male and 25 female) were collected using the Google search engine and were edited using Adobe Photoshop 8.0. Faces were gray-scaled and sized so that the main features fit within a 3.50 deg × 4.81 deg oval (same as Experiment 1). Not all of the faces were familiar to each of the subjects; to establish individual baseline recognition of the 50 celebrities, subjects fixated on the upright faces and identified (named) the famous person in a pretest. Each subject recognized a subset of the 50 faces; average recognition was 34.6/50 (69.2%). Following this pretest, all 50 famous faces (familiar ones and non-familiar ones)

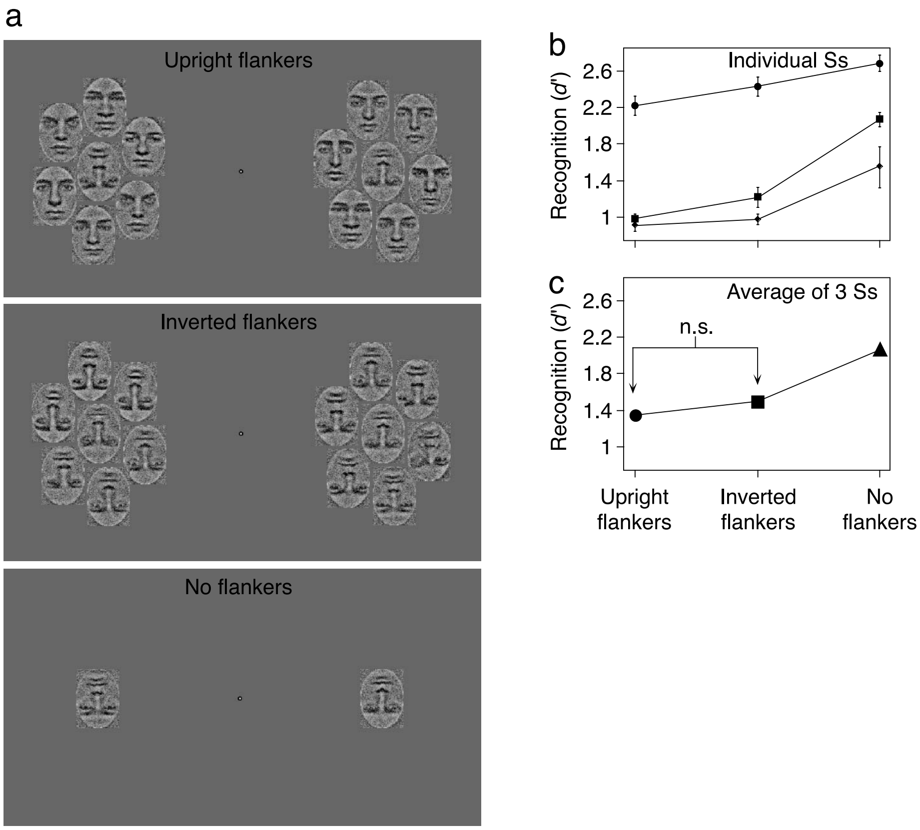

Figure 3. The methods used in Experiment 3 remained the same as those in Experiment 1, except the central target face was inverted (a). Discrimination was impaired with upright and inverted flankers for each of the three subjects, indicating that crowding was effective (b; each symbol represents a separate subject). Panel c shows the average recognition across subjects. Recognition of the target face was better in the inverted flankers condition compared to the upright flankers condition ($F(1, 2) = 11.1$, $p = 0.08$, $\text{eta}^2 = 0.84$; Wilcoxon test, $Z = 1.61$, $p = 0.055$), contrary to what one would expect if the similarity in the orientation of the flankers and target were responsible for the selective crowding effect found in Experiment 1. Error bars, within subjects ±*SEM*.

were then presented in random order at 6.13 deg or 11.58 deg eccentricity for 400 ms, while subjects fixated on the central point and named the famous person. Of the subset of familiar faces (determined in the pretest), the proportion of correctly identified faces was calculated at each eccentricity (i.e., performance was normalized, for each subject, to the set of familiar faces). Because the average number of familiar faces was 34.6, chance performance in the recognition (naming) task was 1/34.6, on average (this is a conservative estimate; chance performance was probably lower than this, considering that subjects did not memorize the set of faces in the pretest).

Results

Target face recognition selectively impaired by upright flankers

In the first experiment, we presented an upright target face either in isolation, or in a crowd of upright or inverted flanker faces (Figures 1a–1c; see Methods). There was a significant difference in recognition between the no flanker and flanker conditions ($F(2, 6) = 14.21$, $p = 0.005$, $\text{eta}^2 = 0.83$), indicating that crowding was effective at impairing recognition (Figures 1d–1f). Importantly, recognition was impaired most when the target face was surrounded by upright compared to inverted flankers (Figure 1f; $F(1, 3) = 22.3$, $p < 0.05$, $\text{eta}^2 = 0.88$; Wilcoxon test, $Z = -1.84$, $p = 0.03$). Not surprisingly, sensitivity decreased significantly with increasing noise added to the images (Figure 1d; $F(4, 12) = 17.92$, $p < 0.05$, $\text{eta}^2 = 0.86$). The results show that target face identification was significantly and selectively impaired in the upright flankers condition.

Target house recognition not selectively impaired by upright flankers

In the second experiment, we tested whether the same crowding effect would occur with objects that are not holistically processed, using images of houses as stimuli (Figure 2a). Unlike the results predicted for Experiment 1, we did not expect to find a significant difference between the two flanker conditions (upright vs. inverted) because houses are processed in a part-based manner, regardless of orientation. Figures 2b–2c show that the presence of flankers effectively disrupted recognition of target houses (crowding condition effect, $F(2, 6) = 9.86$, $p = 0.013$, $\text{eta}^2 = 0.77$). However, there was not a significant difference in recognition of the target house when it was surrounded by either upright or inverted flanker houses ($F(1, 3) = 0.075$, $p > 0.05$, $\text{eta}^2 = 0.02$). Thus, there is crowding of house targets by house flankers, but this crowding is not gated by orientation.

Impaired target face recognition not due to flanker similarity

Could the selective impairment in Experiment 1 be due to the similarity of the upright flanker faces interfering with recognition of the upright target face? To address this question, we used inverted faces as the central target stimuli in Experiment 3. If similarity were responsible for the impaired performance in Experiment 1, we would expect observers' performance to be worse in the inverted flankers condition and *better* in the upright flankers condition. However, this was not the case. Figures 3b and 3c show that recognition of inverted targets was worse in the upright flankers condition compared to the inverted flankers condition ($F(1, 2) = 11.1$, $p = 0.08$, $\text{eta}^2 = 0.84$; Wilcoxon test, $Z = 1.61$, $p = 0.055$). Note that the trend of this effect was opposite from that predicted by the similarity argument, and each subject showed the same pattern of results (Figure 3b), demonstrating that a floor effect cannot be responsible for the results. More importantly, when compared to the first experiment, there was a significant interaction between the flanker–target similarity and the recognition of the target (Figure 4a; $F(1, 2) = 34.7$, $p < 0.05$, $\text{eta}^2 = 0.95$). The interaction was also significant between the first and second experiments (Figure 4b; $F(1, 3) = 9.9$, $p < 0.05$, $\text{eta}^2 = 0.76$). Finally, we analyzed the correct trials (hits)—independent of the constant proportion of false alarms (Figure 4c)—and found the same pattern of results and the same significant interactions (Figure 4d), showing that the impaired recognition in the first experiment was unique to upright target faces with upright flankers.

To more closely examine whether the similarity in the orientation of the flanker and target stimuli can explain the results in Experiment 1, each subject's data in the two flanker conditions (upright vs. inverted) in all three experiments were directly compared. According to the similarity argument, recognition should be most impaired when the target and flankers have the same orientation. To test this prediction, we subtracted the discrimination values in the similar flanker condition from those in the dissimilar flanker condition for each experiment, within each subject (Figure 5a). When the data are normalized in this way, positive scores along the ordinate (Figure 5) would indicate impaired recognition when the flankers had an orientation similar to that of the target. If orientation similarity were responsible for the crowding effect in Experiment 1, then we should observe positive scores in Figure 5 across all three experiments (dashed line). Experiments 2 and 3, however, did not show impaired discrimination for similarly oriented flankers,

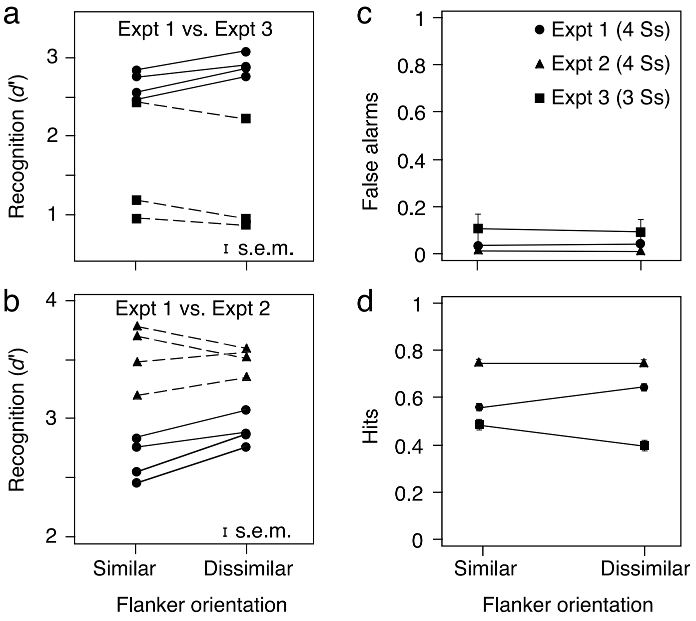

Figure 4. Comparison of the two crowding conditions in each of the first three experiments. Along the abscissa is the similarity in the orientation of the flankers relative to the target. Panel a illustrates that there was a significant interaction between recognition performance and the target–flanker similarity across the first (circles) and third (squares) experiments ($F(1, 2) = 34.7$, $p < 0.05$, $\text{eta}^2 = 0.95$). That is, in the first experiment, recognition was most impaired when an upright target face was surrounded by similarly oriented flankers (upright faces). In the third experiment, on the other hand, inverted target face recognition was most impaired when surrounded by dissimilarly oriented flankers (upright faces), indicated by the dashed line. Comparing the first (circles) and second (triangles) experiments (b) also revealed a significant interaction ($F(1, 3) = 9.9$, $p < 0.05$, $\text{eta}^2 = 0.76$). Error bars in the bottom right corner of the graphs are representative within-subjects *SEM*. The proportion of false alarms (trials in which a target was incorrectly reported as present) across the two flanker conditions was very small (c), indicating that subjects employed strict criterions; the fact that the false alarm rate was constant across the different conditions validates direct *d′* comparisons (MacMillan & Creelman, 2004). A selective analysis of the correct responses (hits) revealed the same pattern of results (d), and the same significant interactions between the first and third experiments ($F(1, 2) = 114.7$, $p < 0.05$, $\text{eta}^2 = 0.98$) and the first and second experiments ($F(1, 3) = 18.8$, $p < 0.05$, $\text{eta}^2 = 0.86$). Within-subject error bars (±*SEM*) are smaller than some symbols.

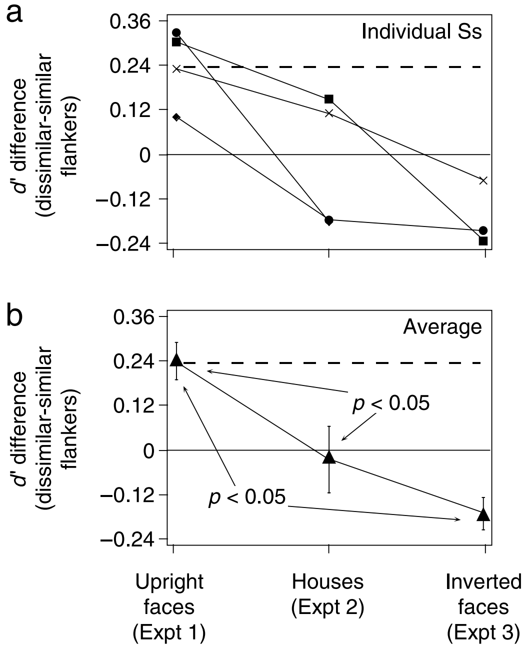

Figure 5. Summary of the flanker conditions in all three experiments for each subject showing that orientation similarity was not responsible for the results of Experiment 1. According to the similarity prediction, when the target and flankers have similar orientations, recognition will be impaired. For each of the three experiments (abscissa), the ordinate shows the normalized recognition (the difference in *d*-prime for dissimilar minus similar flankers). In each of the three experiments, the orientation of the flankers could be similar or dissimilar to the central target; in the first two experiments, upright flankers were similar to the upright targets, and in the third experiment, inverted flankers were similar to the inverted targets. Positive values on the ordinate indicate that similar flankers reduced recognition (e.g., in the first experiment, the upright flankers were more effective at impairing discrimination). The dashed line indicates the prediction of similarity—the expected results if the similar orientation of the targets and flankers was responsible for the crowding effect. The data do not obey the similarity prediction. Panel B shows that there is a significant difference between the three experiments ($F(2, 8) = 8.6$, $p = 0.01$, eta^2 = 0.68). Planned post hoc comparisons revealed that the first experiment was significantly different than both the second ($t(6) = 2.8$, two-tailed $p < 0.05$, $r = 0.75$) and third ($t(5) = 5.5$, two-tailed $p < 0.05$, $r = 0.93$) experiments. There was not a significant difference between the second and third control experiments ($t(5) = 1.26$, two-tailed $p = 0.26$). The results show that the similarity between the orientation of targets and flankers is not responsible for the impaired recognition in the first experiment. Error bars, between subjects ±*SEM*.

and this deviation from the similarity prediction was significant ($F(2, 8) = 8.6$, $p = 0.01$, eta^2 = 0.68; a non-parametric Kruskal–Wallis test on the individual subject data confirmed this significant interaction, $\chi^2(2) = 9.8$, $p < 0.05$). Planned post hoc comparisons revealed that there was a significant difference in the normalized discrimination score for the first versus the second experiment ($t(6) = 2.8$, two-tailed $p < 0.05$, $r = 0.75$), and for the first versus the third experiment ($t(5) = 5.5$, two-tailed $p < 0.05$, $r = 0.93$). There was not a significant difference between the second and third experiments ($t(5) = 1.26$, two-tailed $p = 0.26$). The results of this analysis provide strong evidence that similarity was not responsible for the selective crowding found in Experiment 1.

Selective interference between upright faces is crowding, not masking

The fourth experiment tested the eccentricity dependence of the crowding effect found in Experiment 1. Figure 6 shows that there were main effects of eccentricity ($F(2, 4) = 8.9$, $p = 0.034$, eta^2 = 0.82) and crowding ($F(2, 4) = 22.1$, $p = 0.007$, eta^2 = 0.92). The eccentricity effect indicates that faces are harder (though not impossible) to recognize at eccentric locations, consistent with previous findings (Goren & Wilson, 2006; McKone, 2004). More importantly, the crowding effect increased significantly with increasing stimulus eccentricity (crowding × eccentricity interaction; $F(4, 8) = 9.9$, $p = 0.003$). At the fovea, there was no significant crowding ($F(2, 4) = 3.2$, $p > 0.05$). The eccentricity dependence of the crowding and the fact that it did not occur at the fovea suggest that masking is not responsible for the results (Pelli et al., 2004). To ensure that the lack of a crowding effect at the fovea was not due to ceiling performance, two subjects from Experiment 5 completed 6 additional runs with more noise added to the central faces (Figure 6, insets). The open and solid gray symbols in Figure 6 show that increasing the noise reduced overall accuracy, but there was still no difference in performance across the three flanker conditions ($F(2, 2) = 0.77$, $p = 0.56$). This indicates that the lack of a crowding effect observed at the fovea was not due to subjects' ceiling performance.

Consistent with the first experiment, recognition of targets presented at 11.58 deg eccentricity was impaired by the presence of flankers (Figure 6). Paired *t*-tests revealed a significant difference between the upright and inverted flanker conditions ($t(5) = 4.6$, $p = 0.006$), and the upright and no flanker conditions ($t(5) = 4.5$, $p = 0.007$). The eccentricity dependence of these results is consistent with previous definitions of crowding (Bouma, 1970; Pelli et al., 2004) and demonstrates that the impaired face recognition due to upright flankers observed in Experiment 1 is a result of crowding, and not masking or some other effect.

Famous face recognition at eccentric locations

The purpose of Experiment 5 was to confirm that subjects can recognize familiar faces presented at eccentric locations, and to rule out the possibility that subjects might rely on simple feature detection when faces are

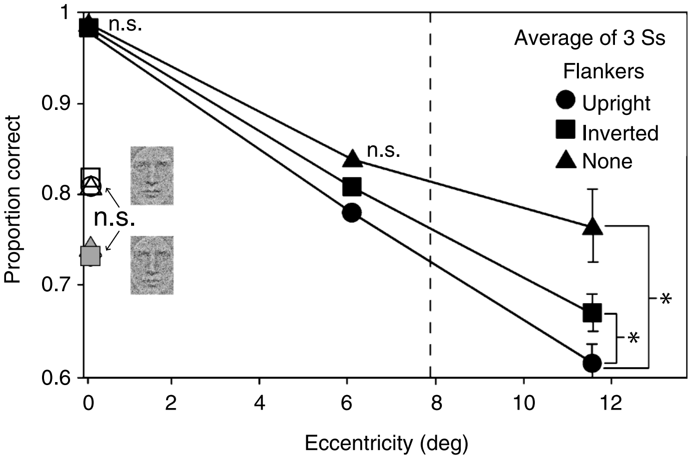

Figure 6. Holistic crowding as a function of stimulus eccentricity. The holistic crowding effect-selective crowding of upright target faces by upright flankers (Figure 1) was modulated by the eccentricity of the stimuli. All three subjects performed above chance at all eccentricities (chance level was 33% correct). As the stimuli were presented more foveally (abscissa), the crowding effect decreased ($F(4, 8) = 9.9$, $p = 0.003$, $\text{eta}^2 = 0.83$). Pairwise comparisons revealed that, at 11.58 deg eccentricity, upright flankers impaired recognition significantly more than inverted flankers ($t(5) = 4.6$, $p = 0.006$) and significantly more than no flankers ($t(5) = 4.5$, $p = 0.007$). At 0 deg and 6.13 deg eccentricity, there was not a significant difference between the upright and inverted flanker conditions (the more significant effect was at 6.13 deg; $t(2) = 1.9$, $p = 0.10$ one-tailed). To ensure that the lack of a crowding effect at 0 deg eccentricity was not due to subjects' ceiling performance, we tested two subjects in 6 additional runs with more noise added to the central faces (open and solid gray symbols show performance at two noise levels, see inset face images). With increasing noise, overall accuracy was reduced, but there was no differential crowding effect ($F(2, 2) = 0.77$, $p = 0.56$). Asterisks indicate significant pairwise comparisons ($p < 0.05$). In all conditions, the center-to-center separation between the target and flankers was 3.94 deg. The dashed vertical line indicates the eccentricity that was twice this separation and is the approximate point at which crowding occurs in experiments using other kinds of features and letters (Bouma, 1970; Pelli et al., 2004). The data closely follow this half-eccentricity rule, supporting the conclusion that the selective impairment of upright face recognition by upright face flankers at eccentric locations is a genuine crowding effect, and not due to pattern masking or salience of upright faces. Error bars, between subjects ±*SEM*.

presented in the periphery. Consistent with previous studies (Goren & Wilson, 2006; McKone, 2004), subjects were significantly above chance at recognizing famous faces at all tested eccentricities (least significant condition was inverted faces at 11.58 deg, $t(6) = 5.6$, $p = 0.001$). Figure 7 shows that there were significant effects of eccentricity ($F(1, 6) = 10.6$, $p = 0.018$) and inversion ($F(1, 6) = 10.9$, $p = 0.017$) on face recognition. The significant effect of eccentricity supports previous findings that face recognition becomes difficult in the periphery (Goren & Wilson, 2006; McKone, 2004), and that this is likely due to self-crowding of the features within the face (Martelli et al., 2005). Nevertheless, peripheral face recognition is far from impossible (Figure 7). Further, the fact that there was a significant inversion effect (difference between upright and inverted face recognition) at 11.58 deg eccentricity ($t(6) = 8.4$, $p = 0.001$) indicates that subjects were not just performing a simple detection task of a single feature but were using configural or holistic information to identify the faces (Farah et al., 1995; Freire, Lee, & Symons, 2000; Leder & Bruce, 2000).

Discussion

The experiments reported here suggest that, in addition to low-level crowding, there is selective crowding

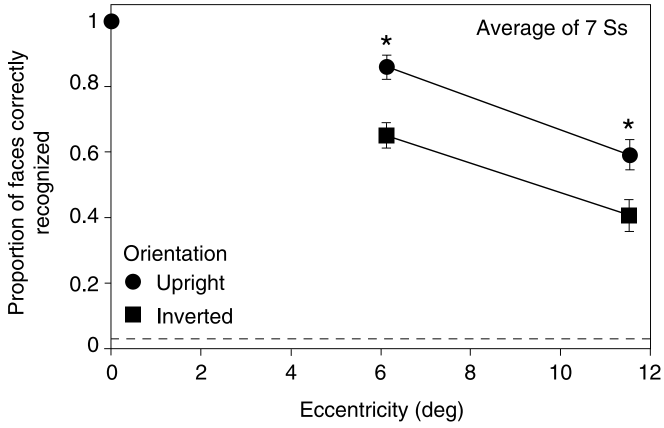

Figure 7. Eccentric recognition of famous faces without crowding. Famous faces were presented in isolation at one of three eccentricities (foveal, 6.13 deg, or 11.58 deg, identical to the eccentricities of the faces in the previous experiments). Subjects first identified (named) 50 famous faces while fixating the face. For each subject, the subset of recognized faces (mean = 69% correct, 34.6 faces) were then presented in a random order at one of two eccentricities (abscissa). Subjects were required to identify (name) the famous face presented in their periphery. The ordinate shows the proportion of correctly identified faces (normalized to the total number of faces that each subject recognized; hence, the data point at 100% correct at 0 deg eccentricity). Average chance level performance was 2.9% (1/34.6), represented by the dashed line. At all eccentricities, subjects were significantly above chance performance. There were significant effects of eccentricity and inversion ($F(1, 4) = 13.1$, $p < 0.05$, and $F(1, 4) = 9.4$, $p < 0.05$, respectively). Inverting the faces impaired recognition at each eccentricity (least significant pairwise comparison was in the 6.13 deg eccentricity condition, $t(6) = 3.1$, $p < 0.05$). The results demonstrate that subjects are able to identify familiar faces at eccentric locations, and this recognition is not restricted to feature-based processes. Error bars, within subjects ±*SEM*.

between high-level representations of faces. The results of Experiment 1 demonstrated that when an upright target face was surrounded by upright flanker faces, recognition was significantly worse than when the target was surrounded by inverted flankers or none at all. However, this occurred only when the central stimuli were upright *faces* and cannot be attributed to similarity. Further, the crowding effect displayed eccentricity dependence and did not occur at the fovea, ruling out alternative explanations.

In addition to the crowding between upright representations of faces, Experiments 2 and 3 revealed crowding at a lower level of analysis. In Experiment 2 (house recognition), subjects' average *d*-prime scores were nearly identical for upright and inverted house flankers. This is not surprising because house images are processed in a part-based manner; because this feature-based processing strategy is employed for both upright and inverted houses, we would not expect a difference in performance. Experiment 3 showed that inverted target faces were more impaired by upright than inverted flanker faces, though this effect was not significant. Still, this is an interesting result and could be due to the fact that the upright face flankers carry feature and configural information, both of which might crowd the inverted target. The inverted flankers, on the other hand, can only exert feature-level crowding. The fact that the inverted target face is processed by features does not preclude the possibility that nearby configural information (at a super-ordinate level) trumps or crowds the feature level information of the inverted target.

How do we know that the selective interference between upright faces is crowding and not salience, masking, or some other effect? We addressed this in Experiment 4 by measuring crowding as a function of eccentricity. If the salience or distractibility of upright face flankers was responsible for the results in Figure 1, then we should have observed an interference effect at all eccentricities—even at the fovea (Palermo & Rhodes, 2002). However, we found that the selective interference between upright faces only occurred when the target–flanker separation was less than half the eccentricity of the target. Moreover, there was no interference found at the fovea, which rules out upright face salience and masking as explanations. Together, these results demonstrate a precise spatial dependence of the effect we observed in Experiment 1, and they are consistent with the operational definition of crowding proposed by Pelli et al. (2004).

The results here suggest that there is crowding between configural (upright) representations of faces. This is just one type of crowding at one particular level; however, it does not discount past research that has characterized other levels of crowding. Figures 6 and 7 clearly illustrate that crowding also occurs within a single face (Martelli et al., 2005): Recognition of an isolated face declined significantly with increasing eccentricity. This supports Martelli and colleagues' (2005) conclusion that crowding among the features within a face, or self-crowding, is one of the critical limits to face recognition in the periphery. Indeed, self-crowding may be the single most influential limitation on peripheral face recognition (Figures 6 and 7). This is not the only type of crowding, however, and it cannot explain our results. For example, the feature-to-feature separations (e.g., bars, edges, facial features) were constant in all conditions, and yet we still observed selective crowding between upright faces. Therefore, in addition to feature-level crowding within the face, our results demonstrate that crowding occurs selectively between the holistic or configural representations of upright faces as well.

The results here are somewhat surprising, given what is known about the neural mechanisms of face recognition. Our results show spatially precise interactions between holistic representations of faces, which would seem to require a region that topographically codes configural information about faces. The configural or holistic information about faces is believed to be analyzed in the fusiform face area (Schiltz & Rossion, 2006; Yovel & Kanwisher, 2005). The FFA, however, has been reported as either non-retinotopic or coarsely retinotopic (Levy, Hasson, Avidan, Hendler, & Malach, 2001; Malach, Levy, & Hasson, 2002). The lack of fine retinotopy in the FFA, however, does not mean that it codes images with position invariance. Even with very large receptive fields, if there is sufficient overlap across the population of neurons, the FFA could effectively carry a coarse-code for face position on a very precise scale (Eurich & Schwegler, 1997). The fact that FFA topography results are currently mixed is less revealing about FFA architecture than it is about the limitations of current fMRI analytic techniques; with advances in fMRI, a more precise picture of FFA topography will soon emerge. Even if the FFA is not ultimately responsible, our results indicate that some other region or network must carry both configural information about faces and precise spatial information.

Whether due to mandatory grouping, averaging, interference, lateral masking, or attention, current models of crowding (Ariely, 2001; Blake, Tadin, Sobel, Raissian, & Chong, 2006; Chung et al., 2001; He et al., 1997; Intriligator & Cavanagh, 2001; Levi et al., 1985; Parkes et al., 2001; Pelli et al., 2004) must be updated to account for the fact that high-level representations of objects can selectively crowd each other. Likewise, models that posit a single mechanism for crowding or suggest that it operates at a single level in the visual system must be revised in light of the likelihood that crowding occurs at multiple stages of processing.

Acknowledgments

Thanks to Ken Nakayama and Frank Tong for providing stimuli, to Thuc-Nhi Nguyen, Tom Harp, Jennifer Temple, Lica Iwaki, Nicole Spotswood, Jason Fischer, Santani Teng, and Nicole Wurnitsch for help collecting data, and to Denis Pelli, Marialuisa Martelli, and an anonymous reviewer for helpful comments on an earlier draft. This work was supported by NIH.

Commercial relationships: none.

Corresponding author: David Whitney.
Email: dwhitney@ucdavis.edu.
Address: Center for Mind and Brain, and Department of Psychology, UC Davis, One Shields Avenue, Davis, CA 95618, USA.

References

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

Blake, R., Tadin, D., Sobel, K. V., Raissian, T. A., & Chong, S. C. (2006). Strength of early visual adaptation depends on visual awareness. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 4783–4788. [PubMed] [Article]

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

Boutet, I., & Chaudhuri, A. (2001). Multistability of overlapped face stimuli is dependent upon orientation. *Perception, 30,* 743–753. [PubMed]

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 335,* 95–103. [PubMed] [Article]

Chung, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research, 41,* 1833–1850. [PubMed]

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107–117. [PubMed]

Eurich, C., & Schwegler, H. (1997). Coarse coding: Calculation of the resolution achieved by a population of large receptive field neurons. *Biological Cybernetics, 76,* 357–363. [PubMed]

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21,* 628–634. [PubMed]

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105,* 482–498. [PubMed]

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research, 33,* 173–193. [PubMed]

Freire, A., Lee, K., & Symons, L. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29,* 159–170. [PubMed]

Goren, D., & Wilson, H. R. (2006). Quantifying facial expression recognition across viewing conditions. *Vision Research, 46,* 1253–1262. [PubMed]

Green, D. M., Weber, D. L., & Duncan, J. E. (1977). Detection and recognition of pure tones in noise. *Journal of the Acoustical Society of America, 62,* 948–954. [PubMed]

Haase, S. J., & Fisk, G. (2001). Confidence in word detection predicts word identification: Implications for an unconscious perception paradigm. *American Journal of Psychology, 114,* 439–468. [PubMed]

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383,* 334–337. [PubMed]

He, S., Cavanagh, P., & Intriligator, J. (1997). Attentional resolution. *Trends in Cognitive Sciences, 1,* 115–121.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43,* 171–216. [PubMed]

Latham, K., & Whitaker, D. (1996). Relative roles of resolution and spatial interference in foveal and peripheral vision. *Ophthalmic & Physiological Optics, 16,* 49–57. [PubMed]

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53,* 513–536. [PubMed]

Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research, 25,* 963–977. [PubMed]

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience, 4,* 533–539. [PubMed] [Article]

MacMillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences, 6,* 176–184. [PubMed]

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://journalofvision.org/5/1/6/, doi:10.1167/5.1.6. [PubMed] [Article]

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6,* 255–260. [PubMed]

McKone, E. (2004). Isolating the special component of face recognition: Peripheral identification and a Mooney face. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 181–197. [PubMed]

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 573–599. [PubMed]

Palermo, R., & Rhodes, G. (2002). The influence of divided attention on holistic face perception. *Cognition, 82,* 225–257. [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,* 739–744. [PubMed] [Article]

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://journalofvision.org/4/12/12/, doi:10.1167/4.12.12. [PubMed] [Article]

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces? *Cognition, 88,* 79–107. [PubMed]

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition, 103,* 34–79. [PubMed]

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage, 32,* 1385–1394. [PubMed]

Starr, S. J., Metz, C. E., Lusted, L. B., & Goodenough, D. J. (1975). Visual detection and localization of radiographic images. *Radiology, 116,* 533–538. [PubMed]

Strasburger, H., Harvey, L. O., Jr., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception & Psychophysics, 49,* 495–508. [PubMed]

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: A Human Experimental Psychology, 46,* 225–245. [PubMed]

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9,* 483–484. [PubMed]

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

Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with Vernier acuity. *Vision Research, 15,* 1137–1141. [PubMed]

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology: Human Perception and Performance, 81,* 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16,* 747–759. [PubMed]

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology, 15,* 2256–2262. [PubMed] [Article]