Are faces processed like words? A diagnostic test for recognition by parts

Marialuisa Martelli
Psychology and Neural Science, New York University
New York, NY, USA
Fondazione Santa Lucia, I.R.C.C.S
Rome, Italy

Najib J. Majaj
Psychology and Neural Science, New York University
New York, NY, USA

Denis G. Pelli
Psychology and Neural Science, New York University
New York, NY, USA

Correspondence should be addressed to M. Martelli. e-mail: mlm9@nyu.edu

Keywords: face recognition, word recognition, feature integration, crowding, isolation, recognition by parts, holistic, inversion, face superiority

Do we identify an object as a whole or by its parts? This deceptively simple question has been surprisingly hard to answer. It has been suggested that faces are recognized as wholes and words are recognized by parts. This paper extends the observation of crowding between objects to crowding between the parts of an object. We applied a diagnostic test for crowding to words and faces, and we find that the critical spacing of the parts required for recognition is proportional to distance from fixation and independent of size and kind. The critical spacing of crowding defines an isolation field, a region that isolates the target. Some objects can only be recognized when the parts are isolated from each other by the critical spacing. In that case, recognition is by parts. Recognition is holistic if the observer can recognize the object even when the whole object fits within a critical spacing. To assess the robustness of the crowding test we manipulated familiarity, i.e. whole-object superiority and inversion. We find that threshold contrast for word and face identification is the product of two factors: familiarity and crowding. Familiarity increases sensitivity by a factor of ×1.5, independent of eccentricity, while crowding attenuates sensitivity more and more as eccentricity increases. Our findings show that observers process words and faces in much the same way: the effects of familiarity and crowding do not distinguish between them. Words and faces are both recognized by parts, and their parts – letters and facial features – are recognized holistically. We propose that internal crowding be taken as the signature of recognition by-parts.

Introduction

Psychophysical proposals for how people recognize objects have largely been bottom-up, building on what is known about feature detection. Cognitive proposals have been top-down, reasoning from what is known about object categorization.

Object identification begins with independent feature detection and then proceeds to integration (Neisser, 1967; Campbell & Robson, 1968; Robson & Graham, 1981; Pelli, Farell, and Moore, 2003). A feature is an independently detected component of the image, much less than a letter. Modern psychophysics mainly focuses on the problem of how we integrate features to recognize the object. Gestalt psychologists noted that we seem to recognize objects holistically; the perceived shape is not simply the sum of the parts (Wertheimer, 1923). This idea stimulated investigation of how we represent objects. The contemporary debate focuses on whether we recognize particular objects holistically or by-parts (Prinzmetal, 1995). However, attempts to empirically distinguish between these computations have had only limited success (see Rakover, 2002, for an overview).

According to several cognitive models, we recognize objects through a hierarchical process that includes a part-based stage (e.g. Marr & Nishihara, 1978; Johnston & McClelland, 1980; Biederman 1987). It has been proposed that good object parts are nameable, functional, and are segments defined by extrema of concave curvature (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Tversky & Hemenway, 1984; Hoffman & Richards, 1984; Diamond and Carey, 1986). In a face, good parts are the eyes, nose, and mouth; in a word, the letters (Farah, Wilson, Drain & Tanaka, 1998).

It has been suggested that many objects are recognized by-parts, but that faces are recognized primarily as wholes (Farah, 1991; Farah, 1995; Farah, Wilson, Drain & Tanaka, 1998). The face superiority and inversion effects are perhaps the best existing evidence for the holistic encoding of faces (Valentine, 1988; Farah, Tanaka & Drain, 1995). In the face *superiority* effect, observers better discriminate a part when presented in the context of a face than when presented alone or in a scrambled face (Tanaka & Farah, 1993; Tanaka & Sengco, 1997). In the *inversion* effect, a face is harder to recognize when upside-down (Yin, 1969; Farah, Tanaka & Drain, 1995; but see Sekuler, Gaspar, Gold & Bennett, 2004). Superiority of the face over a face part is interpreted as evidence for a holistic process that deals simultaneously with the entire pattern: when a face part is isolated from the rest, performance worsens. However, words too show a superiority effect: it is easier to identify a letter when presented in a word context than when presented in isolation or in a nonword context (Reicher, 1969; Wheeler, 1970), but the effect is too small to reject the hypothesis that word recognition is strictly letter or feature based (Pelli, Farell & Moore, 2003).

If faces and words are processed differently, then they might be analyzed by different modules in the brain (Fodor, 1983; Biederman, 1987; Ullman, 1989; Tarr & Buelthoff, 1998). In a groundbreaking review of the pattern of co-occurrence of impairments of face, object, and word recognition in a large group of brain-damaged patients, Farah (1991) boldly suggested that the brain has separate modules for different kinds of object, with faces and words falling at opposite ends of a shape processing continuum: Faces are processed as wholes and words are processed by parts.

While there is controversy over how holistic face recognition might be implemented (Smith, 1967; Diamond & Carey, 1986; Schyns, 1998; Farah, Wilson, Drain & Tanaka, 1998; Gauthier, Behrmann & Tarr, 1999; Wenger & Ingvalson, 2002), fMRI studies have revived the idea that faces and words are processed in separate modules. Several studies have succeeded in isolating face-specific regions in the brain that seem anatomically distinct from regions selective for buildings, letters, words, and body parts (Kanwisher, McDermott & Chun, 1997; Aguirre, Zarahn & D'Esposito, 1998; Polk & Farah, 1998; Kanwisher, Stanley & Harris, 1999; Downing, Jiang, Shuman & Kanwisher, 2001; Grill-Spector, Kourtzi & Kanwisher, 2001).

Another approach to understanding the difference in processing between faces and other objects considers the development through childhood of face recognition. Even as neonates, humans prefer looking at faces over other objects, which suggests that an innate component of face recognition may contribute to development of the face area (Goren, Sarty & Wu, 1975; Johnson, Dziurawiec, Ellis, and Morton, 1991). However, it has also been proposed that the face area is really an expertise area, and that faces are special only because we are so practiced and competent in judging them (Diamond & Carey, 1986; Gauthier & Tarr, 1997; Gauthier, Skudlarski, Gore & Anderson, 2000).

In this paper we look for crowding in faces and words as a symptom of recognition by parts. Crowding describes the impairment of recognizability of an object by neighboring objects. Unlike ordinary masking, in which the object disappears, in crowding the object remains visible but it is unidentifiable. Ordinary masking affects feature detection while crowding impairs feature integration. Crowding has mostly been measured between letters. When the flanker letters are close to the target letter, the target remains visible but its features are jumbled with those of the flankers. It's a big effect. Observers "see" jumbled shapes that are hard to describe. Object identification becomes easy when the flankers are moved far enough away from the target. Critical spacing is how far away the flankers must be to allow recognition of the target. In ordinary masking, critical spacing is proportional to size, independent of eccentricity. In crowding, critical spacing increases with eccentricity. It is roughly half of the viewing eccentricity, independent of target and flanker size (Bouma, 1970; Strasburger, Harvey & Rentschler, 1991). This is the diagnostic test for crowding (Pelli, Palomares, and Majaj, 2004).

Crowding is known as interference between objects. Here we examine crowding between the parts of an object. If its parts crowd each other then the object crowds itself and is unrecognizable. Indeed, when the object is a word the letters crowd each other, and the word is unrecognizable (Bouma, 1973). We applied the crowding test to parts of faces and words. The critical spacing of crowding defines an "isolation field", a region at the target location over which the observer integrates features to compute any multi-feature object property demanded by the task (Pelli, Palomares, and Majaj, 2004). Critical spacing defines how much of the visual field must be isolated in order for the object to be recognized. Earlier authors have used various names for a region over which features are integrated as revealed by crowding experiments (Levi, Klein & Aitsebaomo, 1985; Toet & Levi, 1992; Hess, Dakin & Kapoor, 2000; Intriligator & Cavanagh, 2001; Parkes, Lund, Angelucci, Solomon & Morgan, 2001).

Recognition is *by-parts* when the observer can recognize the object only if the parts are moved apart so as to isolate them from each other by the critical spacing. Recognition is *holistic* when observers can recognize the object even if the extent of the whole object is less than a critical spacing. As noted above, the critical spacing is roughly half the eccentricity (Bouma, 1970).

Here we address whether faces and words are processed differently: holistically or by parts. We take the parts of a word to be letters. We take the parts of a face to be the mouth, nose, eyes, hair, and outline. We present faces and words at various eccentricities and we vary the spacing between the parts to measure critical spacing. If the object is recognized holistically it can be recognized even when the whole object lies within a critical spacing, without isolating any part. If recognition is by parts then object identification will be possible only when the parts are isolated from each other by the critical spacing. We manipulate familiarity to address the robustness of our diagnosis. We present faces and words in a familiar (right-side up) and in an unfamiliar arrangement (nonwords and upside-down words and faces).

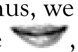
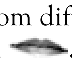
Exp. 1 measures face and word recognition as a function of eccentricity and finds an inferiority effect that grows with eccentricity. Exp. 2 addresses whether the word and face inferiority effects are due to crowding and whether word and face parts interact in the same way. In Exp. 3 we look for a difference between faces and words in the familiarity effect. The results decompose the effect of context into two factors: effects of familiarity and crowding.

Methods

Observers

Seven observers with normal or corrected-to-normal vision participated. One observer (MM) is an author. The other observers were paid by the hour. TA, AS, AB, and MM observed faces. TG, MS, and HS observed letters. All observers completed a 2,000-trial learning phase prior to collecting the data reported here.

Stimuli

As face stimuli we used both photoes and caricatures. A face and three mouth pictures were selected from the Paul Ekman face photo database (http://www.paulekman.com). The database contains the facial expressions of the basic emotions (Ekman, 1992). We built part and part-in-context stimuli, using the mouth as the part. Martelli, Majaj, Palomares, Leigh, Ekman & Pelli (2001) show that when the face parts are easily discriminable (i.e. presence or absence of the teeth) observers don't show a face superiority effect. Thus, we selected three mouths from different faces: smile , frown , and neutral . As context, we selected a female face from the same set of photos. A face and three mouth caricatures were selected from the Lar DeSouza database (http://www.lartist.com/celebrity.htm). We presented the mouth alone and in the context of the caricature of a female face. We selected three mouths from the database: thin , fat  and medium . In separate runs the mouths were presented alone or in context, right-side-up or upside-down. Observers were asked to identify the mouth.

We used an alphabet of five letters cgprx rendered in the Bookman font by Adobe Type Manager. We designed the word context to be uninformative as to the target letter identity (e.g. ace, age, ape, are, axe). In each run we used several word contexts. Nonwords were built by equating the identity of the first and last letter (e.g. aca, aga, ara, axa). Combinations that generated words, acronyms, or abbreviations were discarded (e.g. apa). For words and nonwords, the target was always the central letter. In separate runs we presented the letters alone or in the word or nonword context. We also presented letters and words both right-side-up and rotated 180 deg. Observers were asked to identify the target letter.

When the signal size was fixed (Exps. 1 and 3) the mouth size was 1.5 deg and the letter size was 0.8 deg. Mouth size is measured horizontally end to end. Letter size is "x height", i.e. the height of the letter x.

Procedure

In each trial the target was a random sample from the signal set. The set included 3 signals in the case of face photos and caricatures, and 5 signals in the case of words. Each signal presentation was accompanied by a beep. A response screen followed, showing all the possible signals at 80% contrast. One of the signals in the response screen was otherwise identical to the target. Observers were instructed to identify the signal by clicking on one of the candidates in the response screen. A correct response was rewarded by a beep.

All experiments were performed on Apple Power Macintosh computers using MATLAB software with the Psychophysics Toolbox extensions (http://psychtoolbox.org) (Brainard, 1997; Pelli, 1997). Observers viewed a gamma-corrected grayscale monitor (Pelli & Zhang, 1991) with a background luminance of 16 cd/m^2. The fixation point was a 0.15 deg black square. For foveal viewing, the fixation point was presented for 200 ms. For peripheral viewing, the fixation point appeared and remained on the screen for the entire duration of the trial. In either case, 400 ms after the fixation point appeared, the signal appeared for 200 ms. The signal was always presented in the center of the screen. The viewing eccentricity of the signal was set by positioning the fixation point on the screen. With faces, the fixation point was positioned at the same height as the center of the mouth. With words, the fixation point was positioned at half of the letter "x height". For peripheral viewing, the stimuli were always presented in the right visual field.

When face photos were used, either the mouth alone or the mouth in context was "pasted" onto a background square with the same average luminance as the face. Letters and caricatures were drawn in white on the gray background. Signal contrast is defined as the ratio of luminance increment to background luminance. When the signal was presented in context, the context received the same contrast reduction as the target part, relative to the original word or face. The observer's threshold contrast was estimated in a 40-trial run, using the improved QUEST staircase procedure using a threshold criterion of 82% correct (Watson & Pelli, 1983; King-Smith, Grigsby, Vingrys, Benes & Supowit, 1994). Log thresholds were averaged over three runs for each condition.

Exp.1 Superiority and inferiority

Exp. 1 measures the whole-object superiority effect as a function of eccentricity for the three kinds of object. Signal size was fixed, the mouth size was 1.5 deg and the letter size was 0.8 deg. For 1.5 deg mouths, efficiency (Pelli & Farell, 1999) of observers AB and TA is independent of viewing eccentricity, E_{ideal}/E^{+} = 8% at 0 and 8 deg. Similarly, Pelli, Burns, Farell & Moore (2004) found that efficiency is the same at 0 and 5 deg eccentricity for letters that are well above the acuity limit, as ours were. We measured contrast threshold for the part alone and in the face or word context at 0, 2, 4, 6, and 8 deg from fixation in the right visual field.

Exp. 2 Crowding

Exp. 2 measures the effect of crowding for words and faces by increasing the part spacing. We used only words and face caricatures because we cannot easily spread out the features in a photo of a face without introducing new features (edges) and destroying old ones.

Part spacing was measured horizontally center-to-center, from letter to letter or from mouth to the nearest facial feature on the horizontal meridian. Toet and Levi (1992) showed that the isolation fields are elliptical, centered on the target, with the main axis oriented toward the fovea. In crowding, threshold contrast for identifying the target part drops with increasing spacing. Critical spacing is the minimum spacing at which there is practically no effect of the flanker. We measure it as the lower break point in a clipped line fit of threshold contrast as a function of spacing (Fig. 4a). Our words and faces were displayed so that crowding extended most horizontally, and we measured critical spacing horizontally.

To test for crowding, we measured critical spacing at 4, 6, 8, and 12 deg eccentricity at one size (0.8 deg letter and 1.5 deg mouth), and at 12 deg eccentricity as a function of size (0.4 to 3.2 deg letter and 0.8 to 3.0 deg mouth). The rest of the parts were proportionally scaled. The facial features never overlapped, even at the smallest spacing.

Exp. 3 Familiarity

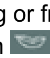
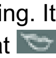
Exp. 3 measures the effect of familiarity as a function of eccentricity for words, face photos, and caricatures. Part size was fixed as in Exp. 1.

We measured threshold contrast for identifying the target part presented in a familiar arrangement (right-side-up words and faces) and in an unfamiliar arrangement (nonwords and upside-down words and faces) at 0, 2, 3, 4, 6, and 8 deg eccentricity. The generation of nonwords is explained above in Stimuli. The familiarity advantage is the ratio of the two thresholds.

Results

Exp. 1. The word and face inferiority effect

We presented the mouth and the central letter alone or in its face or word context. We looked at how context affects recognition across the visual field. Does context help or hinder part recognition? Context helps in the superiority effect, which has often been taken as evidence for holistic processing. Context hinders in crowding between the target part and the rest of the object. If it's crowding, the hindrance with fixed part spacing should grow as eccentricity increases. We measured threshold contrast for identifying the expression of a mouth or a letter with and without the uninformative context of the face or the 3-letter word (Fig. 1). To test for crowding we took our measurements at eccentricities of 0, 2, 4, 6, and 8 deg from fixation. The chosen part size yields equal efficiency of identification of the isolated part in the fovea and periphery (see Methods). As face stimuli, we used both photos and caricatures of faces. Face caricatures produce the same categorical effects as photos (Rhodes, Byatt, Tremewan & Kennedy, 1997; Lewis & Johnston, 1998). We estimated the context advantage by taking the ratio of the observer thresholds for identifying the part (mouth or letter) presented alone and in context (face or word).

Figure 1. **Effect of context: word and face inferiority effect.** (Upper) The word inferiority effect. Fixate on the central square, and try to identify the middle letter on your left. It's hard! Now keep fixating the square and identify the letter on your right. It's easy! The word made it hard to identify the letter. (After Bouma, 1973.) (Middle) The face inferiority effect. Fixate on the central square, and try to tell if the face on the left is smiling or frowning. It's hard! Now keep fixating the square and try to tell if the mouth on the right is smiling or frowning. It's easy! (Lower) Try to tell if the mouth is thin  or fat . Again, it's hard on the left and easy on the right. The face made it hard to identify the shape of the mouth.

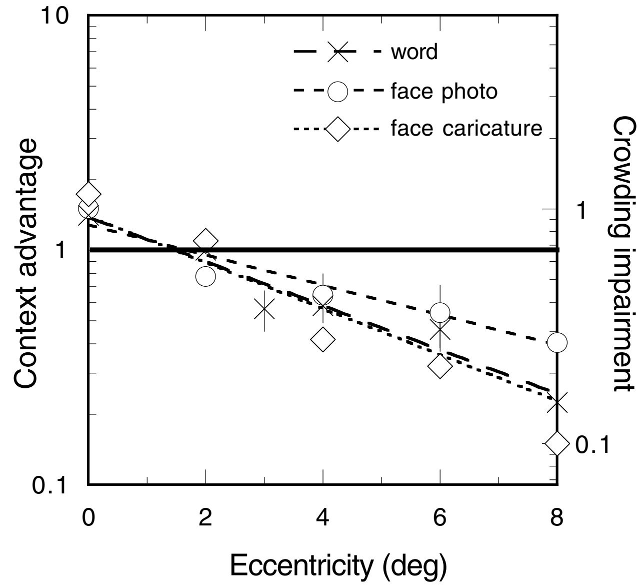


Figure 2. **Effect of context: superiority and inferiority.** (The right vertical scale is explained in Discussion.) Average results for six observers. We measured threshold contrast for identifying the letter or mouth part alone or in the word or face context. We plot the ratio of thresholds for the part alone and in context, averaged across observers, as a function of eccentricity. The part size was fixed, independent of eccentricity (1.5 deg mouth, 0.8 deg letter). The horizontal solid line represents no effect of context. X's are results for words and letters (observers TG, MS, and HS); O's for face and mouth photos (observers AB, TA, and AS); diamonds for face and mouth caricatures (observer MM). Error bars (±1 s.e.) are calculated across observers.

Figure 1 demonstrates the effect. When viewed peripherally, the word or face context hinders identification of the letter or mouth. In central vision, all observers show a superiority effect: they identify a part more easily when it's presented in the context of an uninformative word or face than alone. As Figure 2 shows, it's a small effect, a factor of about 1.6±0.1 in contrast. $m \pm s.e.$ indicates the geometric mean $m = \exp(\text{ave}(\ln(X)))$ and the standard error $s.e. = \text{sqrt}(\text{var}(X)/(n-1))$. Even so, it is an important part of the existing evidence for holistic processing in face recognition. Figure 2 shows the effect. At 0 deg eccentricity it's 1.4±0.1 for words, 1.5±0.1 for face photos, and 1.7±0.1 for face caricatures. In the periphery, we find the opposite — context hinders recognition — and this inferiority effect increases with eccentricity, reaching a factor of 5 for words, 4 for face photos, and 7 for face caricatures at an eccentricity of 8 deg. This is the face and word inferiority effect, whereby, in the periphery, the presence of the face or the word context hinders the observer's identification of the part. The worsening of the inferiority effect with eccentricity is consistent with crowding between the parts and the rest of the object. Our next experiment applies a diagnostic test.

Exp. 2. Crowding

The inferiority effect shows that the face and word context hinders recognition in the periphery. Here we test whether the inferiority effect is due to crowding between the object parts. If crowding occurs between the parts then we should be able to restore recognition by spreading the parts away from each other so that they are isolated by the critical spacing. Alternatively, if crowding occurs between elementary features (e.g. oriented lines), then spacing the parts would not suffice to relieve crowding.

We applied the diagnostic test for crowding to the words and the face caricatures (Fig. 3). We measured threshold contrast for identifying the target part (mouth or central letter) at various eccentricities (0 to 12 deg) and sizes (0.4 to 3.2 deg) as a function of the spacing between the target and the surrounding parts. As illustrated by the first two rows in Fig. 3, for a given target location, we increased spacing by moving the other parts away from the target part. Spacing would also be increased by enlarging the whole face (Fig. 3 third row), but this manipulation confounds size and spacing, so it was not used.

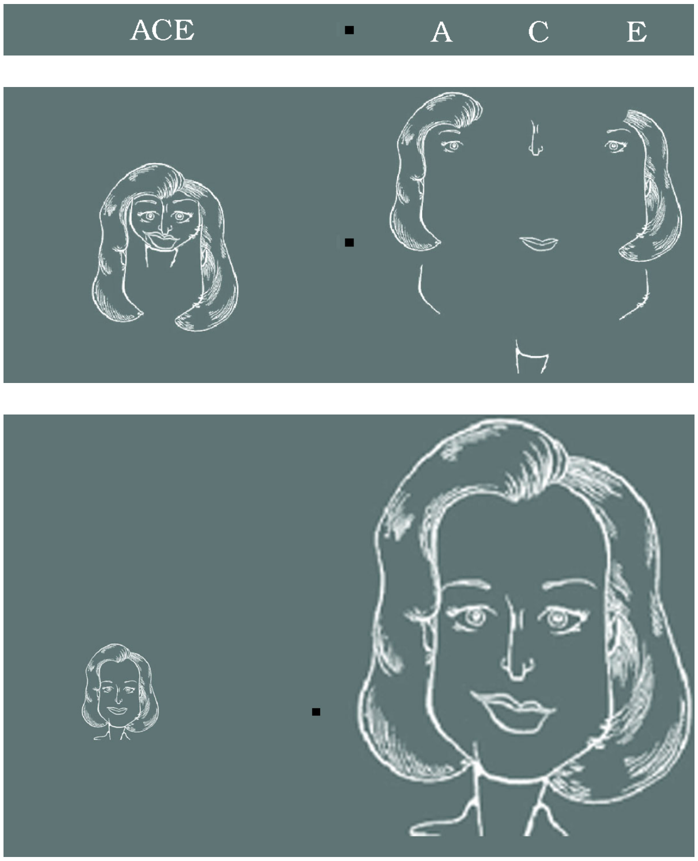
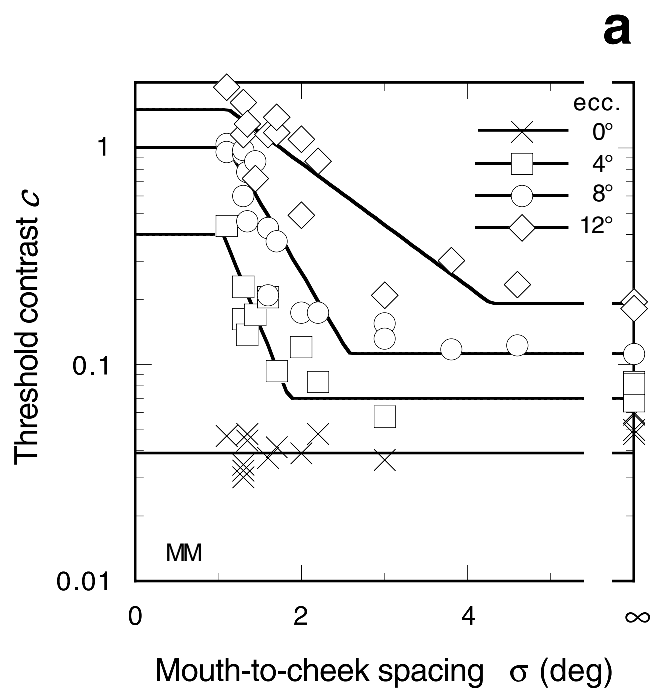


Figure 3. **Measuring critical spacing in words and faces.** In each panel, fixate on the square and try to identify the central letter (C or L?) or mouth (thin or fat ?) on the left and right. As in Fig. 1, it's hard on the left and easy on the right. In the first two panels we increased spacing by moving every part away from the target part, keeping size constant. In the third panel we enlarged the whole face. When the spacing between parts is greater than critical (roughly half of the viewing eccentricity) the other parts don't interfere.

Fig. 4a plots threshold contrast as a function of spacing. For words, we measured the center-to-center horizontal spacing between letters. For faces, we measured the center-to-center spacing between the mouth and the nearest part horizontally. In the fovea, threshold is independent of spacing (horizontal line). The ratio of threshold measured at infinite spacing to that at closer spacing is the face superiority effect. In the periphery, threshold drops with increasing spacing. The results are fit by a clipped line

$$c = \min\left(c_{\text{ceil}}, \max\left(c_{\text{floor}}, ae^{b\sigma}\right)\right) \quad (1)$$

with break points at floor c_{floor} and ceiling c_{ceil} (Pelli, Palomares & Majaj, 2004). The floor break point is critical spacing (Fig. 4a). R^2 of the fit ranged from 0.8 to 0.94.



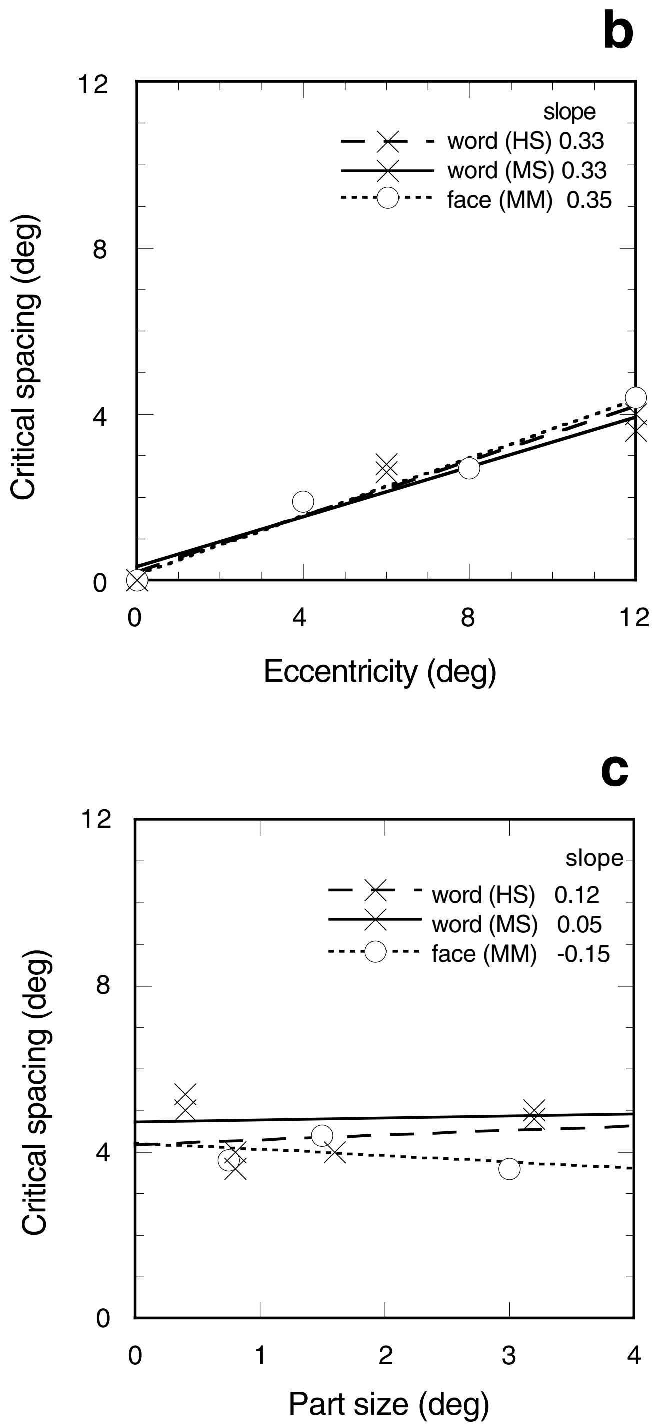


Figure 4. **Diagnostic test for crowding in face caricatures and words.** (a) Threshold contrast as a function of center-to-center part spacing at various eccentricities. The mouth size is 1.5 deg. The results are fit by a clipped line with break points at floor and ceiling. The floor break point is critical spacing. (b) Critical spacing as a function of eccentricity. Critical spacing is proportional to eccentricity with an average slope of 0.34. Letter size is 0.8 deg; mouth size is 1.5 deg. This result is independent of size, as shown in panel c. (c) Critical spacing as a function of part size. Critical spacing is practically independent of part size, with an average slope of 0.007. Eccentricity is 12 deg. The results show that critical spacing is proportional to eccentricity and independent of size. This is the signature of crowding (Pelli, Palomares, and Majaj, 2004). Thus, identifiability of letters and mouths in words and faces in the periphery is limited by crowding between the parts. [JOV: Please lay out Fig. 4 and its caption to appear all together on one page.]

We plot the critical spacing as a function of eccentricity (Fig. 4b) and part size (Fig. 4c). In the fovea, the range of crowding is tiny, only a few minutes of arc (Bouma, 1970), so 1 deg objects like ours would have to overlap in order to crowd, making it difficult to distinguish effects of crowding from ordinary masking, so in plotting Fig. 4b we assume zero critical spacing at 0 deg eccentricity. For all observers, and with both caricatures (X) and words (O), the critical spacing is proportional to viewing eccentricity with an average slope of 0.34, in agreement with Bouma's estimate of roughly 0.5, and R^2 ranging from 0.91 to 0.98 (Fig. 4b). Fig. 4c shows that critical spacing is independent of part size with an average slope of 0.007. We fit a regression line through the data for each observer. R^2 ranges from 0.01 to 0.17. These results show that critical spacing is proportional to eccentricity and independent of size. This is the signature of crowding. In ordinary masking, critical spacing is proportional to size, independent of eccentricity (Pelli, Palomares & Majaj, 2004). Our result that spreading the parts relieves crowding, indicates that face and word recognition requires isolation of the parts. If instead, crowding occurred between elementary features (e.g. oriented lines), then isolating the facial features or the letters would not suffice.

We can calculate the magnitude of the inferiority effect (Exp.1) from the data in Fig. 4a. The threshold elevation (re floor) is the amplitude of the inferiority effect. From Exp. 2 the inferiority effect at 12 deg eccentricity is approximately a factor of 10 for caricatures and a factor of 12 for words.

Exp. 3. Familiarity

Here we measure the familiarity advantage as a function of eccentricity (0 to 8 deg), using the same photos, caricatures, and words as in Exp. 1. The stimuli were presented in familiar (right-side up) and unfamiliar (upside-down faces and words, and nonwords) arrangements. Observers were asked to identify the mouth or the target letter, alone or in context. The part spacing was the same as in Exp. 1, which is well within the critical spacing measured in Exp. 2. As in Exp. 1, the context advantage is the ratio of the thresholds for identifying the part alone and in context. The ratio of the context advantages in the familiar and unfamiliar conditions is the *object familiarity advantage* (Fig. 5). The observers show the same ×1.5±0.1 advantage of familiarity for faces and words, independent of eccentricity. This is consistent with Fine's (2004) finding that the benefit of a word context in reducing the stimulus duration required to identify a letter is independent of eccentricity.

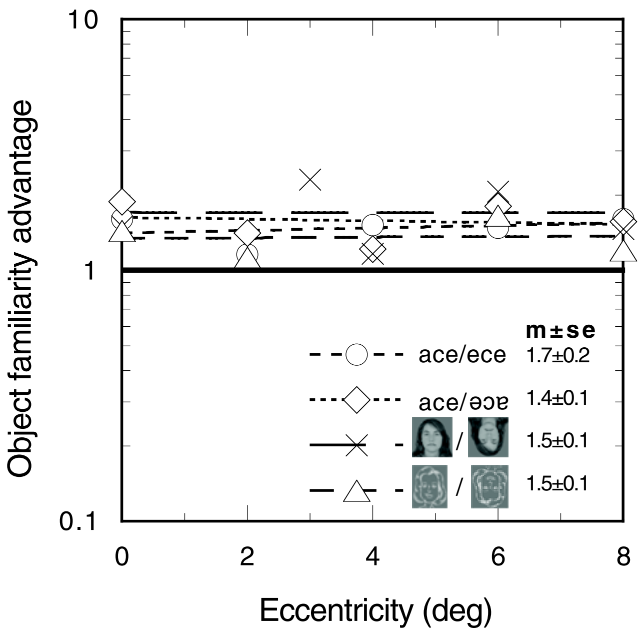


Figure 5. **Familiarity and eccentricity.** We measured threshold contrast for identifying the part alone and in context, in a familiar (right-side-up word or face) and in an unfamiliar arrangement (nonword and upside-down word or face). Context advantage is the ratio of thresholds for the part alone and in context. Object familiarity advantage is the ratio of context advantages obtained in the familiar and unfamiliar conditions. This is plotted for four observers as a function of eccentricity. All the points are above the (solid) equality line. The advantage is the same for words (observers HS, MS) and faces (TA, MM), independent of eccentricity. The regression line slopes (words/nonwords –0.001; words/inverted-words 0.003; face photos 0.02; face caricatures –0.01) are not significantly different from zero.

Discussion

In central vision we find a face and word superiority effect consistent with previous findings (Reicher, 1969; Smith, 1969; Wheeler, 1970; Paap, Newsome, McDonald & Schvaneveldt, 1982; Tanaka & Farah, 1993; Jordan & deBruijn, 1993; Farah, Wilson, Drain & Tanaka, 1998). However, in the periphery we find a much bigger effect in the opposite direction. Threshold contrast is reduced slightly ÷1.5 centrally and increased greatly ×5 at 8 deg in the periphery. The presence of the surrounding face or word helps identification slightly in the central field, and hinders greatly in the periphery. We call this hindrance the *face and word inferiority effect*.

Context both helps and hinders. Exps. 2 and 3 reveal that the context effect is the product of the effects of crowding and familiarity. Eccentricity distinguishes them.

Context hinders through crowding. The key parameter is the spacing between the part (letter or mouth) and the context (rest of the word or face). Letter strings and facial expressions can be identified only if their parts are spaced far enough apart to avoid crowding.

Context helps through familiarity. The familiarity effect is small, increasing contrast sensitivity by a factor of 1.5, independent of eccentricity. (Contrast sensitivity is the reciprocal of threshold contrast.) We can estimate the crowding effect at all eccentricities by dividing out the ×1.5 familiarity effect from the measured context effect in Exp. 1. Thus Fig. 2, using the right scale, plots the crowding effect as a function of eccentricity. The effect of crowding worsens with eccentricity, from ×1 at 0 deg to ×0.17 at 8 deg.

What do these results tell us about object recognition?

Faces are like words

We find the same familiarity advantage for words and faces. We chose words and faces because they have been thought to represent opposite ends of the object spectrum. Words are processed by parts and faces may be processed as wholes (Rumelhart & McClelland, 1982; Farah, 1991; Pelli, Farell & Moore, 2003). Even so, in all our tasks, faces show the same familiarity effects as words do.

We examined two familiarity effects: whole-object superiority and inversion. Neither superiority nor inversion is specific to faces, but inversion is affected by expertise and superiority isn't (Tanaka & Gauthier, 1997; Gauthier, Behrmann & Tarr, 1999). The superiority and inversion effects have the same magnitude (Farah, Wilson, Drain & Tanaka, 1998). However, superiority is learned quickly, in a few hours, and inversion slowly, over many years (Diamond & Carey, 1977; Hay & Cox, 2000; Martelli, Baweja, Mishra, Chen, Fox, Majaj & Pelli, 2002). This difference in learning rate suggests that the two effects are due to different mechanisms, which makes it all the more remarkable that they both affect faces and words equally.

Table 1 surveys the familiarity advantage, including inversion and superiority effects, found by other authors for expert observers of various objects. (Proportions correct have been converted to an equivalent threshold contrast elevation factor, using available estimates of the psychometric function.) These are diverse experiments, so one must be cautious in comparing their results, but it is clear from the table that the inversion and superiority effects have similar magnitudes for words, faces, and other objects such as dogs, landscapes and Greebles. Our familiarity effects for words and faces are identical to theirs. The fact that we obtain the same familiarity effect through inversion and superiority with words and faces undermines the notion that faces are special. By these measures, faces, words, dogs, landscapes, and Greebles are all equally special for expert observers.

We still don't know how people recognize words and faces, but the fact that both tasks show the same effects of crowding and familiarity favors the null hypothesis that faces and words are processed in the same way.

Exp.	Effect	Contrast ratio
Wheeler, 1970	word sup.	1.6
Jordan & deBruijn, 1993	word sup.	1.5
Reicher, 1969	word sup.	1.5
Babkoff, Faust & Lavidor, 1997	word sup.	1.3
Pelli, Farell & Moore, 2003	word sup.	1.3
This study	word sup.	1.5
This study	word inv.	1.4
Tanaka & Sengco, 1997	face sup.	1.6
This study	face sup.	1.6
Tanaka & Farah, 1993	face sup.	1.5
Diamond & Carey, 1986	face inv.	1.9
Tanaka & Sengco, 1997	face inv.	1.9
Yin, 1969	face inv.	1.9
Leder & Bruce, 2000	face inv.	1.8
McKone, Martini & Nakayama, 2001	face inv.	1.8
Sekuler, Gaspar, Gold & Bennett, 2004	face inv.	1.5
This study	face inv.	1.5
Farah, Wilson, Drain & Tanaka, 1995	face inv.	1.4
Farah, Wilson, Drain & Tanaka, 1998	face inv.	1.4
Gauthier, Tarr, Anderson, Skudlarski & Gore, 1999	face inv.	1.4
Tanaka & Farah, 1993	face inv.	1.4
Diamond & Carey, 1986	dog inv.	1.6
Gauthier & Tarr, 1997	Greeble sup.	1.5
Diamond & Carey, 1986	landscape inv.	1.4
Weisstein & Harris, 1974	shape sup.	1.3

Table 1. **The familiarity effect expressed as a contrast ratio.** From each paper we extract the proportion correct p_1 with and p_2 without the familiar context, and estimate the effect context has on threshold contrast. The psychometric function describing how proportion correct for object detection (Nachmias, 1981) and identification (Strasburger, 2001) grows with contrast has a stereotyped shape. For identification this is well described by a Weibull function,

$$p = 1 - (1-\gamma)\exp\left(-(c / c_p)^{\beta}\right)$$

with $\gamma = 1/n$ and $\beta = 1.8$, where c is contrast and c_p is threshold contrast. Solving for c / c_p as a function of p, we calculate the contrast ratio c_{p_2} / c_{p_1} corresponding to proportions correct p_1 and p_2,

$$\frac{c_{p_2}}{c_{p_1}} = \left(\frac{\log \frac{1-p_2}{1-\gamma}}{\log \frac{1-p_1}{1-\gamma}} \right)^{-1/\beta}.$$

A different choice for β will scale all the contrast ratios up or down by a fixed factor. Sekuler et al. (2004) measured threshold contrast for upright and inverted faces, so we simply took the ratio of their thresholds. The magnitude of the familiarity effect is similar for all these objects and tasks. Excluding our results, the geometric mean of the contrast ratio is 1.4±0.1 for words, 1.6±0.1 for faces, and 1.3 for three-dimensional shapes. Experts judging other objects show a similar advantage, 1.5±0.1 (dogs, Greebles, and landscapes).

In some of these experiments, performance is contrast limited, in which case the estimated contrast ratio predicts the effect of familiarity on threshold contrast. Some experiments are not contrast limited. In that case the contrast ratio is merely a transformation, like difference in z-score, that converts two different proportions correct, with and without familiarity, into a single number representing the size of the effect.

Faces and words are recognized by parts

In Farah's 1991 conjecture a face is recognized as a whole and a word by its parts. Are faces and words really recognized so differently? Taking a cognitive top-down approach, we failed to find any difference in the familiarity effect between faces and words. Taking a perceptual bottom-up approach, we measured the region of the visual field that needs to be isolated in order to recognize words and faces, again finding practically identical results for the two kinds of object.

There is abundant evidence that vision detects very simple elementary features (Campbell & Robson, 1968; Robson & Graham, 1981). And there is evidence that we tend to perceive the world as a collection of discrete objects (e.g. Rosch et al., 1976; Di Lollo, Enns & Rensink, 2002). It has been suggested that visual recognition involves an intermediate part-based representation, between elementary features and objects (Biederman, 1987). Proposed object parts in perception include nameable or functional components, and object contours parsed at extrema of concave curvature. In a face, the parts are the eyes, nose, and mouth; in a word, the letters (Farah, Wilson, Drain & Tanaka, 1998). Words and faces are recognized holistically if the visual system goes directly from the elementary features to the whole object representation without recognizing parts.

Crowding has always been described between objects. Here we found crowding within an object. A face or word is unrecognizable in the periphery unless it is huge (Fig. 3). Recognition becomes possible when the parts are spaced far enough apart so that each is isolated from the rest by the critical spacing. Exploding the face or the word (separating the parts as in Exp. 2) isolates the target part (mouth or letter), relieving crowding. This shows that the observer requires isolation of the part for recognition, and that recognizability of the isolated part is essential for recognition of the object.

Not all peripherally viewed objects crowd themselves. Recognition is holistic if it doesn't require isolation of any part. A mouth or letter presented alone in the periphery is recognized. It doesn't crowd itself. It is recognized holistically (Fig. 1).

If the observer required isolation of an elementary feature (e.g. oriented line), then we would expect to have to isolate the elementary feature, and it wouldn't be sufficient for recognition to isolate a facial feature or a letter, because each contains many elementary features in close proximity to one another.

Crowding manipulations reveal how much of the visual field must be isolated for recognition. A word or a face within the critical spacing is unrecognizable. It becomes recognizable when the parts are isolated from each other by the critical spacing. The fact that a part must be isolated from the rest of the object shows that recognition is not holistic. Faces and words are processed by parts.

Are faces foveal?

We find that faces in the periphery crowd themselves, and are thus unrecognizable. If the fusiform face area is more active when the face is recognized then our finding predicts that a face will activate the face area less when presented in the periphery than when presented in the fovea. Levy et al. (2001) isolated face selective regions in the brain and compared activation when a face was presented in the fovea or in the periphery (Levy, Hasson, Avidan, Hendler & Malach, 2001; Malach, Levy & Hasson, 2002). Their largest face was about 14 deg wide including the hair (presented in a 17.5 deg box). At the 16 deg eccentricity they used, the critical spacing is 8 deg (Bouma, 1970), which cannot isolate the facial features of a 14 deg face. We showed one of their large faces at 16 deg eccentricity to three observers (MS, GC, and EH), and asked, "What is it ?" MS said, "I can see hair and features. Their location is face-like. So it is a face, but I cannot tell the gender." GC said, "There are two black structures enclosing something." EH said, "There is something thick and black around some little black lines. The little lines are messy." We increased the size until they could tell the facial expression. The observers could identify the expression only if the face was at least 19 deg wide. Using faces no more than 14 deg wide, Levy et al. found that in all face-selective regions activation was lower in response to a face presented peripherally than centrally. Their results confirm the prediction of crowding: the fusiform face area is less activated when the facial features are closer than the critical spacing.

Conclusion

Words differ qualitatively and are thought to be recognized by parts. Faces differ parametrically and have been thought to be recognized holistically.

Measurements of the effects of spacing, size, and eccentricity on threshold contrast demonstrate two distinct context effects on part identification: familiarity and crowding. Familiarity helps slightly, independent of eccentricity, and crowding hinders greatly, worsening with increasing eccentricity. The effect of context is the product of the two.

This paper extends the observation of crowding between objects to crowding between parts. Internal crowding is the signature of recognition by parts. Internal crowding reveals that to recognize a face or a word observers need to isolate a part. Words and faces are obviously different, yet our results indicate that both are processed by-parts.

Acknowledgments

We thank Susan Carey, Isabel Gauthier, Roberta Daini, Karin James, and Melanie Palomares for helpful discussion. Thanks to Allison Swezey, Corrina Moucheraud, and Michael Su for their careful observations. Thanks to Tracey Berger and Katharine Tillman for editorial advice. Thanks to Lar DeSouza for letting us use and modify his face caricatures. Supported by NIH grant EY04432 to Denis Pelli.

References

Aguirre, G. K., Zarahn, E. & D'Esposito, M. An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron* **21**, 373-83. (1998).

Babkoff, H., Faust, M. & Lavidor, M Lexical decision, visual hemifield and angle of orientation. *Neuropsychologia* **35**, 487-95. (1997).

Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychol Rev,* **94**, 115-147. (1987).

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

Bouma, H. Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, **13**, *767-782* (1973).

Brainard, D.H. (1997). The Psychophysics Toolbox. Spat Vis, 10 (4), 433-436. [PubMed] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9176952

Campbell, F. W. & Robson, J. G. Application of Fourier analysis to the visibility of gratings. *Journal of Physiology,* **197**, 551-566 (1968).

Di Lollo, V., Enns, J. T., Rensink, R. A. Object substitution without reentry? *Journal of Experimental Psychology: General.* **131**, *594-596* (2002).

Diamond, R. & Carey, S. Developmental changes in the representation of faces. *Journal of Experimental Child Psychology* **23**, 1-22. (1977).

Diamond, R. & Carey, S. Why faces are and are not special: an effect of expertise. *J Exp Psychol Gen* **115**, 107-17. (1986).

Downing, P. E., Jiang, Y., Shuman, M. & Kanwisher, N. A cortical area selective for visual processing of the human body. *Science* **293**, 2470-3. (2001).

Ekman, P. Are there basic emotions? *Psychol Rev* **99**, 550-3 (1992).

Farah, M.J. Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology,* **8**, 1-19 (1991).

Farah, M.J. Visual agnosia: disorders of object recognition and what they tell us about normal vision. Cambridge, MA: *MIT press*. (1995)

Farah, M. J., Tanaka, J. W. & Drain, H. M. What causes the face inversion effect? *J Exp Psychol Hum Percept Perform* **21**, 628-34. (1995).

Farah, M. J., Wilson, K. D., Drain, H. M. & Tanaka, J. R. The inverted face inversion effect in prosopagnosia: evidence for mandatory, face-specific perceptual mechanisms. *Vision Res* **35**, 2089-93. (1995).

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

Fine, E.M. (2004). The relative benefit of word context is a constant proportion of letter identification time. *Perception and Psychophysics,* in press [PubMed]

Fodor, J. A. (1983). *The Modularity of Mind.* Cambridge, MA: MIT Press.

Gauthier, I. & Tarr, M. J. Becoming a "Greeble" expert: exploring mechanisms for face recognition. *Vision Res* **37**, 1673-82. (1997).

Gauthier, I., Behrmann, M. & Tarr, M. J. Can face recognition really be dissociated from object recognition? *J Cogn Neurosci* **11**, 349-70. (1999).

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P. & Gore, J. C. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci* **2**, 568-73. (1999).

Gauthier, I., Skudlarski, P., Gore, J. C. & Anderson, A. W. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* **3**, 191-7. (2000).

Grill-Spector, K., Kourtzi, Z. & Kanwisher, N. The lateral occipital complex and its role in object recognition. *Vision Res* **41**, 1409-22 (2001).

Hay,D.C., Cox, R. Developmental changes in the recognition of faces and facial features. *Infant-and-Child-Development* **9**, 199-212. (2000).

Hess, R. F., Dakin, S. C. & Kapoor, N. The foveal 'crowding' effect: physics or physiology? *Vision Res,* **40**, 365-70 (2000).

Hoffman, D. D. & Richards, W. A. Parts of recognition. *Cognition*, **8**, *65-96* (1984).

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognit Psychol, 43 (3),* 171-216. [PubMed] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11689021

Johnson, M. H., Dziurawiec, S., Ellis, H., and Morton, J. (1991) Newborns' preferential tracking of face-like stimuli and its subsequent decline. Cognition, 40 (1-2), 1-19.

Johnston, J. C. and McClelland, J. L. (1980) Experimental tests of a hierarchical model of word identification. *Journal of Verbal Learning and Verbal Behavior, 19 (5),* 503-524

Jordan, T. R. & deBruijn, O. Word superiority over isolated letters: the neglected role of flanking mask contours. *J Exp Psychol Hum Percept Perform* **19**, 549-563. (1993).

Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* **17**, 4302-11. (1997).

Kanwisher, N., Stanley, D. & Harris, A. The fusiform face area is selective for faces not animals. *Neuroreport* **10**, 183-7. (1999).

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C. & Supowit, A.. Efficiency and unbiased modifications of the QUEST threshold method: theory, simulations, Exp.al evaluation and practical implementation. *Vision Research,* **34**, 885-912 (1994).

Leder, H. & Bruce, V. When inverted faces are recognized: the role of configural information in face recognition. *Q J Exp Psychol A* **53**, 513-36. (2000).

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

Levy, I., Hasson, U., Avidan, G., Hendler, T. & Malach, R. Center-periphery organization of human object areas. *Nat Neurosci* **4**, 533-9. (2001).

Lewis, M. B. & Johnston, R. A. Understanding caricatures of faces. *Q J Exp Psychol A* **51**, 321-46. (1998).

Malach, R., Levy, I. & Hasson, U. The topography of high-order human object areas. *Trends Cogn Sci* **6**, 176-184. (2002).

Marr, D. & Nishihara, H. K.. Representation and recognition of spatial organization of three-dimensional shapes. *Proceeding of the Royal Society of London, Series B, Biological Sciences*, **200**, *269-294* (1978).

Martelli, M., Majaj, N., Palomares, M., Leigh, N., Ekman, P. & Pelli, D.G. Which features depend on which faces? [Abstract]. Journal of Vision, 1(3), 289a. (2001).

Martelli, M., Baweja, G., Mishra, A., Chen, I., Fox, J., Majaj, N.J. & Pelli, D.G. How efficiency for identifying objects improves with age. *Perception*: ECVP 2002 abstracts.

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

Nachmias, J. On the psychometric function for contrast detection. *Vision Res* **21**, 215-23 (1981).

Neisser, U. (1967). *Cognitive Psychology.* (New York: Appleton-Century-Crofts.

Paap, K. R., Newsome, S. L., McDonald, J. E. & Schvaneveldt, R. W. An activation--verification model for letter and word recognition: the word-superiority effect. *Psychol Rev* **89**, 573-94. (1982).

Parkes, L., Lund, J., Angelucci, A., Solomon, J.A. & Morgan, M. Compulsory averaging of crowded orientation signals in human vision. *Nat Neurosci,* **4**, 739-744 (2001).

Pelli, D.G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. Vision Res, 31 (7-8), 1337-1350. [PubMed] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1891822

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis, 10 (4), 437-442. [PubMed] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9176953

Pelli, D.G., & Farell, B. (1999). Why use noise? J Opt Soc Am A Opt Image Sci Vis, 16 (3), 647-653. [PubMed] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10069051

Pelli, D.G., Farell, B., & Moore, D.C. (2003). The remarkable inefficiency of word recognition. Nature, 423 (6941), 752-756. [PubMed] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12802334

Pelli, D. G., Burns, C. W., Farell, B. & Moore, D. C., Identifying letters. *Vision Research*. In press. (2004).

Pelli, D. G., Palomares, M. & Majaj, N. J. Crowding is unlike ordinary masking: Distinguishing feature detection and integration. *Journal of Vision,* In press, (2004).

Polk, T.A., Farah, M. J. The neural development and organization of letter recognition: evidence from functional neuroimaging, computational modeling, and behavioral studies. *Proc Natl Acad Sci U S A* **95**, 847-52. (1998)

Prinzmetal, W. Visual feature integration in a world of objects. *Current Directions in Psychological Science*. **4**, *90-94* (1995).

Rakover, S. S. Featural vs. configurational information in faces: a conceptual and empirical analysis. *British Journal of Psychology*, **93**, *1-30* (2002).

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. *J Exp Psychol* **81**, 275-80. (1969).

Rhodes, G., Byatt, G., Tremewan, T. & Kennedy, A. Facial distinctiveness and the power of caricatures. *Perception* **26**, 207-23 (1997).

Robson, J. G. & Graham, N. Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research,* **21**, 409-418 (1981).

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 8 (3), 382-439.

Rumelhart, D. E. & McClelland, J. L. An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychol Rev* **89**, 60-94. (1982).

Schyns, P. G. Diagnostic recognition: task constraints, object information, and their interactions. *Cognition* **67**, 147-79. (1998).

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. Inversion leads to quantitative, not qualitative changes in face processing. *Current Biology* **14**, *391–396* (2004).

Smith, E. E. Effects of familiarity on stimulus recognition and categorization. *J Exp Psychol* **74**, 324-32. (1967).

Smith, E. E. Familiarity of configuration vs discriminability of features in the visual identification of words. *Psychon Sci* **14**, 261-262. (1969). *Spatial Vision,* **10**, 433-436 (1997).

Strasburger, H. (2001). Invariance of the psychometric function for character recognition across the visual field. *Percept Psychophys, 63 (8),* 1356-1376. [PubMed] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11800462

Strasburger, H., Harvey, L., and Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception and Psychophysics, 49*, 495-508.

Tanaka, J. W. & Farah, M. J. Parts and wholes in face recognition. *Q J Exp Psychol A* **46**, 225-45. (1993).

Tanaka, J. W. & Gauthier, I. Expertise in object and face recognition. In Medin, Goldstone & Schyns, Mechanisms of perceptual lerning. Vol **36**, 83-125. (1997).

Tanaka, J. W. & Sengco, J. A. Features and their configuration in face recognition. *Mem Cognit* **25**, 583-92. (1997).

Tarr, M. J. & Buelthoff, H. H. Image-based object recognition in man, monkey and machine. *Cognition*, **67**, 1-20. (1998).

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

Tversky, B. & Hemenway, K., Objects, parts, and categories. *Journal of Experimental Psychology: General.* **113**, *169-193* (1984).

Valentine, T. Upside-down faces: a review of the effect of inversion upon face recognition. *Br J Psychol* **79**, 471-91. (1988).

Watson, A. B. & Pelli, D. G., QUEST. A Bayesian adaptive psychometric method. *Perception and Psychophysics,* **33**, 113-120 (1983).

Weisstein, N. & Harris, C. S. Visual detection of line segments: An object superiority effect. *Science*, **186**, 752-755 (1974).

Wenger, M. J. & Ingvalson, E. M. A decisional component of holistic encoding. *J Exp Psychol: Lear. Mem. & Cogn.* **28**, 872-92. (2002).

Wertheimer, M. (1923) *Psychologische Forschung* **4,** 301-350; translated in Ellis, W. D. (1938) *A Source Book of Gestalt Psychology.* (Harcourt, Brace, and Co., New York).

Wheeler, D. D. Processes in word recognition. *Cognitive Psychol* **1**, 59-85. (1970).

Yin, R.K. Looking at upside-down faces. *J Exp Psychol* 81, 141-45. (1969).