

# 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



Do we identify an object as a whole or by its parts? This simple question has been surprisingly hard to answer. It has been suggested that faces are recognized as wholes and words are recognized by parts. Here we answer the question by applying a test for crowding. In crowding, a target is harder to identify in the presence of nearby flankers. Previous work has described crowding between objects. We show that crowding also occurs between the parts of an object. Such internal crowding severely impairs perception, identification, and fMRI face-area activation. We apply a diagnostic test for crowding to a word and a face, 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 defines an isolation field around the target. Some objects can be recognized only when each part is isolated from the rest of the object 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. Such an object has only one part. Multiple parts within an isolation field will crowd each other and spoil recognition. To assess the robustness of the crowding test, we manipulated familiarity through inversion and the face- and word-superiority effects. 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  $\times 1.5$ , 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.

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

### 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, & Moore, 2003). A feature is an independently detected component of the image, much smaller than a letter. Modern psychophysics 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 (for an overview, see Rakover, 2002).

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). Good object parts are said to be nameable or functional components or object contours parsed at extrema of concave curvature (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Tversky & Hemenway, 1984; Hoffman & Richards, 1984; Diamond & Carey, 1986). Letters are good parts of a word; the facial features – eyes, nose, and mouth – are good parts of a face (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, Wilson, Drain, & Tanaka, 1995; Farah et al., 1998). The face superiority and inversion ef-

facts 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 facial feature if presented in the context of a face than if presented alone or in a scrambled face (Tanaka & Farah, 1993; Tanaka & Sengco, 1997). In the *inversion effect*, a face is harder to recognize when presented upside-down (Yin, 1969; Farah, Tanaka, et al., 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 with the entire pattern as a whole: A face part is harder to identify when the rest of the face is removed. However, words, though not thought to be recognized holistically, also show an object-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 et al., 2003).

Faces and words may be processed in different ways and by different areas of 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 et al., 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 found 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 to looking at 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).

Here we look for crowding in faces and words as a symptom of recognition by parts. Crowding describes the

impairment of recognizability of a target object by neighboring objects. Unlike ordinary masking, which makes the object disappear, in crowding, the object remains visible but is unrecognizable. Ordinary masking impairs 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. Observers "see" jumbled shapes that are hard to describe. Crowding is a big effect: Threshold contrast for identification is raised tenfold. Object identification becomes easy again when the flankers are moved far enough away from the target.

*Critical spacing* is how far away (center to center) each flanker must be to allow recognition of the target. When spacing is smaller than critical, the presence of the flankers makes recognition of the target harder or impossible. Beyond the critical spacing, recognition is unimpaired, and additional spacing provides no further benefit. The critical spacing is the boundary of a region around the target within which flankers impair recognition and outside of which flankers have no effect. In crowding, critical spacing increases with eccentricity. The critical spacing of crowding is roughly half of the viewing eccentricity, independent of target and flanker size (Bouma, 1970; Strasburger, Harvey, & Rentschler, 1991). Proportional dependence of critical spacing on eccentricity, independent of signal size, is diagnostic of crowding; the converse (proportional dependence on size, independent of eccentricity) indicates ordinary masking (Pelli, Palomares, & Majaj, 2004).

Crowding is known as interference between objects. Here we examine crowding between the parts of an object. If an object's parts crowd each other, then the object crowds itself and is unrecognizable. Indeed, when the object is a less-than-huge word in the periphery (i.e., letter spacing less than half the viewing eccentricity), the letters crowd each other, and the word is unreadable (Bouma, 1973). We apply 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, et al., 2004). Critical spacing defines how much of the object must be isolated for the object to be recognized. Note that critical spacing is defined operationally with reference to the center of the flanker, whereas the isolation field is defined theoretically with reference to the center of each elementary feature in the flanker. Presuming that the features are much smaller than the object or part that they make up, we estimate the isolation field diameter to be the same as the critical spacing, roughly half the eccentricity. Earlier authors have used various other names for a region over which features are integrated: "integration field," "perceptive field," "perceptive hypercolumn," "spatial interference zone," "region of selection," and "association field" (Levi,

Klein, & Aitsebaomo, 1985; Toet & Levi, 1992; Latham & Whitaker, 1996; Intriligator & Cavanagh, 2001; Field, Hayes, & Hess, 1993). Each name has its merits, but the old names all emphasize the still-mysterious process occurring within the field – combining features to recognize an object – so they are necessarily vague, whereas the new name, “isolation field,” concretely specifies the exclusion of everything outside the field. To us it seems that the need for isolation is turning out to be a key insight into the computation underlying object recognition, and thus a good basis for naming this exclusionary field.

Some objects can be recognized even when the whole object falls within a critical spacing (i.e., one isolation field). We call this *holistic recognition*. We define a *part for recognition* as a portion of the object that must be isolated for the object to be recognized. An object that can be recognized holistically has only one part (for recognition). An object with more than just a part (for recognition) is recognized only if each part is separated from the rest of the object by the critical spacing. We call this *recognition by parts*. The critical spacing is roughly half the eccentricity (Bouma, 1970).

Here we address whether faces and words are processed differently: holistically versus 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. These are candidate parts for recognition, independent of whether they are “good parts” in any other sense. 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, then it can be identified 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 each part is isolated from the rest of the object by the critical spacing. Work on crowding indicates that the isolation field integrates all elementary features that fall within it. If the true parts for recognition (requiring isolation) are smaller than we supposed, then isolating our gross “parts” will fail to relieve crowding because multiple small parts will still fall within one isolation field and spoil each other’s recognition.

We manipulate familiarity to assess the robustness of our diagnosis. We present faces and words in a familiar (right-side up) and in unfamiliar arrangements (nonwords and upside-down words and faces).

**Experiment 1** measures face and word recognition as a function of eccentricity and finds an inferiority effect that grows with eccentricity. **Experiment 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 **Experiment 3**, we look for a difference between faces and words in the familiarity effect. The results decompose the effect of context into two factors: 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 photos 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 target part. Martelli et al. (2001) show that when the face parts are very easily discriminable (i.e., presence or absence of the teeth), observers do not show a face superiority effect. Thus, we selected three mouths from different faces: smile , neutral , and frown . As context, we selected a female face from the same set of photos. Additionally, 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 , medium , and fat . In separate runs, the mouths were presented alone or in context, right-side-up, or upside-down. Observers were asked to identify the mouth.

In our word testing, 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 of the target letter identity (e.g., ace, age, ape, are, axe). In each run, we used several word contexts. Nonwords had identical first and last letters (e.g., aca, aga, ara, axa). Combinations that generated words or known 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 (**Experiments 1 and 3**), the mouth size was 1.5 deg and the letter size was 0.8 deg. Mouth size is measured horizontally from end to end. *Letter size* is typographic x-height, the height of the lowercase letter x.

### Procedure

In each trial, the target was a random sample from the signal set. The set included three signals in the case of face photos and caricatures, and five 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<sup>2</sup>. The fixation point was a 0.15-deg black square. For central viewing, the fixation point was presented for 200 ms. For peripheral viewing, the fixation point 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 determined by the location of the fixation point. For peripheral viewing, the signal was always presented in the right visual field. 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 above the baseline of the text.

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 with 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.

## Experiment 1: Superiority and inferiority

**Experiment 1** measures the 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, the 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, and Moore (in press) found that efficiency is the same at 0- and 5-deg eccentricity for letters of any size well above the acuity limit, as ours were. We measured threshold contrast 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.

## Experiment 2: Crowding

**Experiment 2** measures the effect of crowding for words and faces by increasing the spacing between parts. We used only words and face caricatures because we cannot easily separate the features in a photograph 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, with the main axis oriented toward the fovea. In crowding, threshold contrast for identifying the target part drops as spacing increases. Critical spacing is the minimum spacing at which there is practically no effect of the flankers on the target. We measured it as the lower break point in a clipped line fit of threshold contrast as a function of spacing (Figure 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 (0.4–3.2-deg letters and 0.8–3.0-deg mouths). The rest of the parts were proportionally scaled. The facial features never overlapped, even at the smallest spacing.

## Experiment 3: Familiarity

**Experiment 3** measures the effect of familiarity as a function of eccentricity for words, face photos, and caricatures. Part size was fixed, as in **Experiment 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 familiarity advantage is the ratio of the two thresholds, familiar to unfamiliar. The generation of nonwords is explained above in **Stimuli**.

# Results

## Experiment 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? In the object superiority effect, which has often been taken as evidence for holistic processing, context helps. In crowding (of the target part by the rest of the object), context hinders. If there is crowding, the hindrance with fixed spacing between parts 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 (Figure 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).

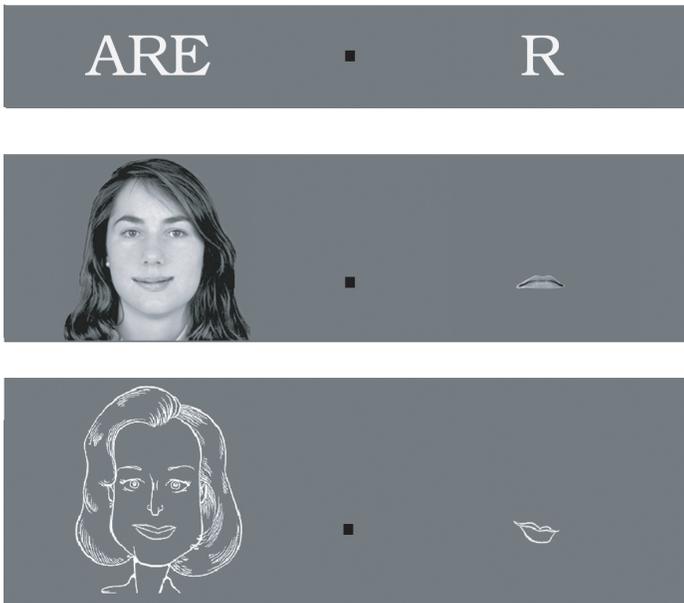


Figure 1. Effect of context: word and face inferiority effect. Upper. The word inferiority effect. Fixate on the central square, and try to identify the middle letter on your left. It's hard! Now keep fixating the square and identify the letter on your right. It's easy! The word made it hard to identify 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 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 an object superiority effect: They identify a part more easily when it is presented in the context of an uninformative word or face than alone. As Figure 2 shows, foveal object superiority is a small effect, a factor of about  $1.6 \pm 0.1$  in contrast.  $M \pm SE$  indicates the geometric mean  $M = \exp(\text{ave}(\ln(X)))$  and the standard error  $SE = \sqrt{\text{var}(X)/(n-1)}$ . Even so, it is an important part of the existing evidence for holistic processing in face recogni-

tion. Figure 2 shows the object superiority effect. The object superiority effect, measured at 0-deg eccentricity, is  $1.4 \pm 0.1$  for words,  $1.5 \pm 0.1$  for face photos, and  $1.7 \pm 0.1$  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 word context hinders the observer's identification of the part. The inferiority effect increases with eccentricity, consistent with the hypothesis that there is crowding between the parts of the object. Our next experiment applies a diagnostic test for crowding.

## Experiment 2: Crowding

The inferiority effect shows that the face and word context hinders recognition of the target part in the periphery. Here we test whether the inferiority effect is due to crowding of the object parts. If the observer must isolate each part to recognize the object, then we should be able to restore recognition by separating each part from the rest of the image by a critical spacing. Alternatively, if the observer must isolate each elementary feature (e.g., oriented lines), then to release recognition from crowding it would be necessary to separate these elementary features from each other, and separating the parts would not suffice to relieve crowding.

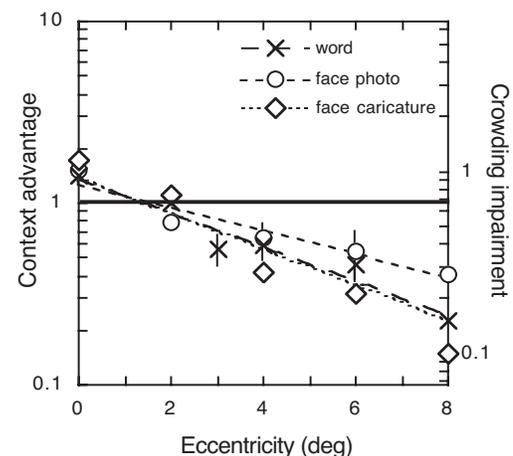


Figure 2. Effect of context: superiority and inferiority. (The right vertical scale is explained in Discussion.) Average results for six observers. We measured threshold contrast for identifying the letter or mouth part alone or in the word or face context. We plot the ratio of thresholds for the part 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.  $\times$ s are results for words and letters (observers TG, MS, and HS); Os are for face and mouth photos (observers AB, TA, and AS); and diamonds are for face and mouth caricatures (observer MM). Error bars ( $\pm 1 SE$ ) are calculated across observers.

We applied the diagnostic test for crowding to the words and the face caricatures (Figure 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 upper two panels of Figure 3, for a given target location, we increased spacing by moving the other parts away from the target part. Spacing could also be increased by enlarging the whole face (Figure 3, bottom panel), but this manipulation confounds size and spacing, so it was not used. However, in their size-scaling study, Mäkelä, Näsänen, Rovamo, and Melmoth (2001) measured threshold contrast for face identification as a function of size at several eccentricities, and we include their results in our analysis below.

Figure 4a plots threshold contrast as a function of spacing. For words, we measured the center-to-center horizontal

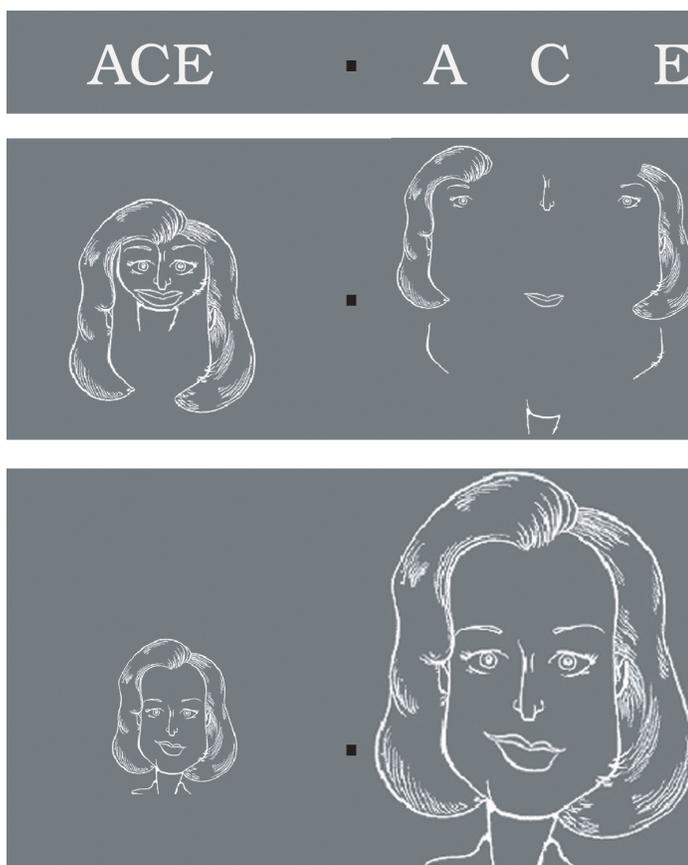


Figure 3. Measuring critical spacing in words and faces. In each panel, fixate on the square and try to identify the central letter (C or L?) or mouth (thin  or fat  ?) on the left and right. As in Figure 1, it is hard on the left and easy on the right. In the first two panels, we increased spacing by moving every other 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 the critical spacing (roughly half of the viewing eccentricity), the other parts do not interfere.

spacing between letters. For faces, we measured the center-to-center spacing between the mouth and the part nearest to it horizontally. At zero eccentricity, threshold is independent of spacing (horizontal line). At zero eccentricity, 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

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

as a function of spacing  $\sigma$ , with break points at floor  $c_{\text{floor}}$  and ceiling  $c_{\text{ceil}}$  (Pelli, Palomares, et al., 2004). The floor break point is the critical spacing, the point where recognition is no longer impaired by crowding, beyond which further spacing provides no additional benefit (Figure 4a).  $R^2$  of the fit ranged from 0.8 to 0.94.

We plot the critical spacing as a function of eccentricity (Figure 4b) and part size (Figure 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 to crowd, making it difficult to distinguish effects of crowding from ordinary masking, so, in plotting Figure 4b, we assume zero critical spacing at 0-deg eccentricity. For all observers, for both caricatures (O) and words (x), Figure 4b shows that 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, with  $R^2$  ranging from 0.91 to 0.98. This is consistent with the size-scaling results of Mäkelä et al. (2001). They measured threshold contrast for face identification as a function of face size at various eccentricities (0 to 10 deg). Plotting the critical spacing estimated from their results as gray diamonds in Figure 4b above shows a similar proportionality with eccentricity. The proportionality constant is lower in their results, presumably because their task (identifying the face) was easier than ours (identifying the mouth). Figure 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 (Pelli, Palomares, et al., 2004). In ordinary masking, critical spacing is proportional to size, independent of eccentricity. Finding that separating 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 to restore recognition.

The amplitude of the inferiority effect is the threshold elevation in Figure 4a. It shows that the inferiority effect at 12-deg eccentricity is big: approximately  $\times 10$  for caricatures and  $\times 12$  for words.

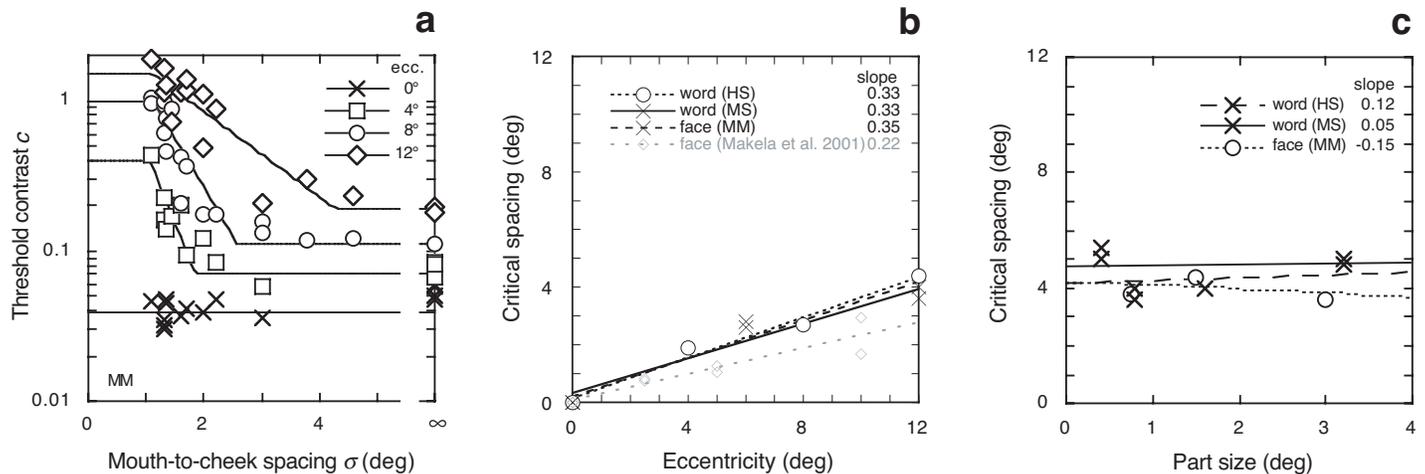


Figure 4. Diagnostic test for crowding in face caricatures and words. (a). Threshold contrast as a function of center-to-center part spacing at various eccentricities. The mouth size is 1.5 deg. The results are fit by a clipped line with break points at floor and ceiling. The floor break point is critical 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. The gray diamonds are based on the threshold contrasts for face identification measured by Mäkelä et al. (2001). We estimated critical size at each eccentricity in their Figures 2A and 2B. We estimate the spacing of facial features (eyes, nose, and mouth) to be 42% of the face size (width of photo in their Figure 1) so critical spacing is 42% of critical size. (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, et al., 2004). Thus, identifiability of letters and mouths in words and faces in the periphery is limited by crowding between the parts.

### Experiment 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 Experiment 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 Experiment 1, well within the critical spacing measured in Experiment 2. As in Experiment 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* (Figure 5). The observers show the same  $\times 1.5 \pm 0.1$  advantage of familiarity for faces and words, independent of eccentricity. This is consistent with Fine's (2004) finding that the benefit of word context in reducing the stimulus duration required to identify a letter is independent of eccentricity.

## 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 & deBrujin, 1993; Farah et al., 1998). However, in the periphery, we find a much bigger effect in the opposite direc-

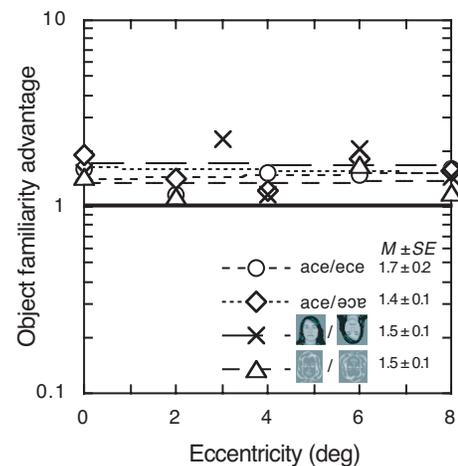


Figure 5. Familiarity and eccentricity. We measured threshold contrast for identifying the part alone and in context, in a familiar (right-side-up word or face) and in an unfamiliar arrangement (nonword and upside-down word or face). Context advantage is the ratio of thresholds for the part alone and 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 and MS) and faces (TA and 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.

tion. Threshold contrast is reduced slightly  $\div 1.5$  centrally and increased greatly  $\times 5$  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. [Experiments 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). Letters and facial features can be identified only if they 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  $\times 1.5$  familiarity effect from the measured context effect in [Experiment 1](#). Thus [Figure 2](#), using the right vertical scale, plots the crowding effect as a function of eccentricity. Crowding worsens as eccentricity increases, from  $\times 1$  at 0 deg to  $\times 0.17$  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 differ qualitatively and are thought to be recognized by parts; faces differ parametrically and have been thought to be recognized holistically (Rumelhart & McClelland, 1982; Farah, 1991; Pelli et al., 2003). Even so, in all our tasks, faces show the same familiarity effects that words do.

We examined two familiarity effects: object superiority and inversion. Neither effect is specific to faces, but inversion is affected by expertise while the object superiority effect is not (Tanaka & Gauthier, 1997; Gauthier, Behrmann & Tarr, 1999). The object-superiority and inversion effects have the same magnitude (Farah et al., 1998). However, the object superiority effect is acquired quickly, in a few hours, and inversion slowly, over many years (Diamond & Carey, 1977; Hay & Cox, 2000; Martelli et al., 2002). This difference in learning rate suggests that the two effects are due to different mechanisms, making it all the more remarkable that they both affect faces and words equally.

[Table 1](#) surveys the familiarity advantage (including the inversion and object 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 object 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. Our finding that words and faces show the same effect of familiarity (inversion and object superiority) 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.

### **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 how much of the object must be isolated to recognize faces and words, 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 parts are 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 intermediate 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 ([Figure 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 [Experiment 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.

We defined a *part for recognition* as a portion of the object that must be isolated for the object to be recognized. For words and faces, we conjectured that the parts for recognition might be letters and facial features. One could imagine a part for recognition to be smaller than we supposed, perhaps a single elementary feature. However, two aspects of our results reject the possibility of smaller parts for recognition. First, if the observer required isolation of smaller parts (e.g., oriented lines), then we would have to separate these smaller parts. It would not be enough to

Experiment	Effect	Contrast ratio
Wheeler, 1970	word sup.	1.6
Jordan & deBrujin, 1993	word sup.	1.5
Reicher, 1969	word sup.	1.5
Babkoff, Faust & Lavidor, 1997	word sup.	1.3
Pelli et al., 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, et al., 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 study, 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,

$$p = 1 - (1 - \gamma)\exp(-(c/c_p)^\beta). \quad (2)$$

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  $p_2$ ,

$$\frac{c_{p_2}}{c_{p_1}} = \left( \frac{\log \frac{1-p_2}{1-\gamma}}{\log \frac{1-p_1}{1-\gamma}} \right)^{-1/\beta}. \quad (3)$$

A different choice for  $\beta$  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  $1.4 \pm 0.1$  for words,  $1.6 \pm 0.1$  for faces, and 1.3 for three-dimensional shapes. Experts judging other objects show a similar advantage,  $1.5 \pm 0.1$  (dogs, Greebles, and landscapes). Note: 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 the difference in z-score, that converts two different proportions correct, with and without familiarity, into a single number representing the size of the effect.

separate the facial features or letters, because each would contain several small parts, which would crowd each other. Second, if letters and facial features are not recognized as units and instead are composed of parts, then, when presented alone, they should be recognized by parts. Instead, we find that they are recognized holistically, the whole con

tained in a single isolation field (Figure 1). Crowding worsens with eccentricity, but efficiency for a letter (Pelli, Burns, et al., in press) and a mouth (see Methods: Experiment 1) of fixed size is independent of eccentricity out to 8 deg from fixation. Thus, letters and mouths do not crowd themselves. They are recognized holistically.

Crowding manipulations reveal how much of the object must be isolated to achieve unimpaired recognition. A word or a face within the critical spacing is unrecognizable. It becomes recognizable when each part is isolated from the rest by the critical spacing. The fact that a part must be isolated from the rest of the object shows that recognition is not holistic. Mouths and letters are recognized holistically, and faces and words are recognized by parts.

### Face area less activated by a crowded face

Unless they are huge, 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 these psychophysical findings predict that a face will activate the face area less when presented peripherally than when presented centrally. Levy, Hasson, Avidan, Hendler, and Malach (2001) identified face-selective regions in the brain and compared activation when a face was presented centrally or peripherally (Malach, Levy, & Hasson, 2002). Presented in a 17.5-deg box, their largest face was 14-deg wide, including the hair. In their 14-deg face, the facial features are about 5 deg apart (from the center of the mouth to the center of the hair measured horizontally), which is less than the critical spacing of roughly 8 deg at the 16-deg eccentricity they used (Bouma, 1970). 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." Using faces no more than 14-deg wide, Levy et al. (2001) 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 face area is less activated when the facial features are closer than the critical spacing.

## Conclusion

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 study extends the observation of crowding between objects to crowding between parts of an object. Internal crowding is the hallmark of recognition by parts. Internal crowding greatly affects subjective report, objective identification, and fMRI face-area activation.

Words differ qualitatively and are thought to be recognized by parts. Faces differ parametrically and have been thought to be recognized holistically. Internal crowding reveals that to recognize a face or a word observers must

isolate a part. Words and faces are obviously different, yet our results indicate that both are recognized by parts.

## Acknowledgments

This is the second in a series of papers about crowding and its cure, isolating to recognize (#1 Pelli, Palomares, et al., 2004; #3 Su, Berger, Majaj, & Pelli, 2004). We thank Diana Balmori, Tracey Berger, Susan Carey, Roberta Daini, Isabel Gauthier, Karin James, Melanie Palomares, Jamie Radner, and Katharine Tillman for helpful discussion. Questions from the reviewers, Martha Farah and anonymous, helped considerably in sharpening the argument. Thanks to Allison Swezey, Corrina Moucheraud, and Michael Su for their careful observations. Thanks to Lar DeSouza for letting us use and modify his face caricatures. Supported by National Institutes of Health Grant EY04432 to DP.

Commercial relationships: none.

Corresponding author: Marialuisa Martelli.

Email: mlm9@nyu.edu.

Address: Department of Psychology, University of Rome La Sapienza, Via dei Marsi 78, 00184, Roma, Italy.

## References

- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: Evidence and implications. *Neuron*, *21*(2), 373-383. [PubMed]
- Babkoff, H., Faust, M., & Lavidor, M. (1997). Lexical decision, visual hemifield and angle of orientation. *Neuropsychologia*, *35*(4), 487-495. [PubMed]
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, *94*(2), 115-147. [PubMed]
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(241), 177-178. [PubMed]
- Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, *13*(4), 767-782. [PubMed]
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433-436. [PubMed]
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the modulation response of the eye. *Journal of Physiology*, *197*, 551-556.
- Di Lollo, V., Enns, J. T., Rensink, R. A. (2002). Object substitution without reentry? *Journal of Experimental Psychology: General*, *131*, 594-596.
- Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology*, *23*(1), 1-22. [PubMed]

- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115(2), 107-117. [PubMed]
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473. [PubMed]
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3), 550-553. [PubMed]
- Farah, M. J. (1991). Patterns of co-occurrence among the associative agnosias: Implications for visual object representation. *Cognitive Neuropsychology*, 8(1), 1-19.
- 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(3), 628-634. [PubMed]
- Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision Research*, 35(14), 2089-2093. [PubMed]
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review*, 105(3), 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(2), 173-193. [PubMed]
- Fine, E. M. (2004). The relative benefit of word context is a constant proportion of letter identification time. *Perception and Psychophysics*, 66(6), 897-907. [PubMed]
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *Journal of Cognitive Neuroscience*, 11(4), 349-370. [PubMed]
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197. [PubMed]
- Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, 37(12), 1673-1682. [PubMed]
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2(6), 568-573. [PubMed]
- Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, 56(4), 544-549. [PubMed]
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10-11), 1409-1422. [PubMed]
- Hay, D. C., & Cox, R. (2000). Developmental changes in the recognition of faces and facial features. *Infant and Child Development*, 9(4), 199-212.
- Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, 18(1-3), 65-96. [PubMed]
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, 43(3), 171-216. [PubMed]
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1-2), 1-19. [PubMed]
- Johnston, J. C., & 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., & deBrujin, O. (1993). Word superiority over isolated letters: The neglected role of flanking mask contours. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 549-563.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302-4311. [PubMed]
- Kanwisher, N., Stanley, D., & Harris, A. (1999). The fusiform face area is selective for faces not animals. *Neuroreport*, 10(1), 183-187. [PubMed]
- King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, 34(7), 885-912. [PubMed]
- Latham, K., & Whitaker, D. (1996). Relative roles of resolution and spatial interference in foveal and peripheral vision. *Ophthalmic and 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*, 53(2), 513-536. [PubMed]
- Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision Research*, 25(7), 963-977. [PubMed]
- Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, 4(5), 533-539. [PubMed]

- Lewis, M. B., & Johnston, R. A. (1998). Understanding caricatures of faces. *Quarterly Journal of Experimental Psychology A*, 51(2), 321-346. [PubMed]
- Mäkelä, P., Näsänen, R., Rovamo, J., & Melmoth, D. (2001). Identification of facial images in peripheral vision. *Vision Research*, 41(5), 599-610. [PubMed]
- Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Science*, 6(4), 176-184. [PubMed]
- Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B*, 200(1140), 269-294. [PubMed]
- Martelli, M., Baweja, G., Mishra, A., Chen, I., Fox, J., Majaj, N. J. & Pelli, D. G. (2002). How efficiency for identifying objects improves with age [Abstract]. *Perception*, 31, ECV Abstracts.
- Martelli, M., Majaj, N., Palomares, M., Leigh, N., Ekman, P., & Pelli, D. G. (2001). Which features depend on which faces? [Abstract] *Journal of Vision*, 1(3), 289a, <http://journalofvision.org/1/3/289/>, doi:10.1167/1.3.289.
- 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(3), 573-599. [PubMed]
- Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research*, 21(2), 215-223. [PubMed]
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: the word-superiority effect. *Psychological Review*, 89(5), 573-594. [PubMed]
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437-442. [PubMed]
- Pelli, D. G., Burns, C. W., Farell, B. & Moore, D. C. (in press). Identifying letters. *Vision Research*.
- Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America A*, 16(3), 647-653. [PubMed]
- Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, 423(6941), 752-756. [PubMed]
- Pelli, D. G., Palomares, M. & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, 4(12), 1136-1169, <http://journalofvision.org/4/12/12/>, doi: 10.1167/4.12.12. [PubMed][Article]
- Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, 31(7-8), 1337-1350. [PubMed]
- Polk, T. A., & Farah, M. J. (1998). The neural development and organization of letter recognition: Evidence from functional neuroimaging, computational modeling, and behavioral studies. *Proceedings of the National Academy of Sciences U. S. A.*, 95(3), 847-852. [PubMed] [Article]
- Prinzmetal, W. (1995). Visual feature integration in a world of objects. *Current Directions in Psychological Science*, 4(3), 90-94.
- Rakover, S. S. (2002). Featural vs. configurational information in faces: A conceptual and empirical analysis. *British Journal of Psychology*, 93(Pt 1), 1-30. [PubMed]
- Reicher, G. M. (1969). Perceptual recognition as a function of the meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275-280. [PubMed]
- Rhodes, G., Byatt, G., Tremewan, T., & Kennedy, A. (1997). Facial distinctiveness and the power of caricatures. *Perception*, 26(2), 207-223. [PubMed]
- Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21(3), 409-418. [PubMed]
- 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. (1982). An interactive activation model of context effects in letter perception. Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89(1), 60-94. [PubMed]
- Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. *Cognition*, 67(1-2), 147-179. [PubMed]
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, 14(5), 391-396. [PubMed]
- Smith, E. E. (1967). Effects of familiarity on stimulus recognition and categorization. *Journal of Experimental Psychology*, 74(3), 324-332. [PubMed]
- Smith, E. E. (1969). Familiarity of configuration vs. discriminability of features in the visual identification of words. *Psychonomic Science*, 14, 261-262.
- Strasburger, H. (2001). Invariance of the psychometric function for character recognition across the visual field. *Perception and Psychophysics*, 63(8), 1356-1376. [PubMed]

- Strasburger, H., Harvey, L. O., Jr., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception and Psychophysics*, 49(6), 495-508. [PubMed]
- Su, M., Berger, T. D., Majaj, N., & Pelli, D. G. (2004). Crowding, shuffling, and capitalizing reveal three processes in reading. Manuscript submitted for publication.
- Tanaka, J., & Gauthier, I. (1997). Expertise in object and face recognition. In R. L. Goldstone (Ed.), *Perceptual learning: The psychology of learning and motivation*, Vol. 36 (pp. 83-125). San Diego, CA: Academic Press.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A*, 46(2), 225-245. [PubMed]
- Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, 25(5), 583-592. [PubMed]
- Tarr, M. J., & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, 67(1-2), 1-20. [PubMed]
- Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, 32(7), 1349-1357. [PubMed]
- Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113(2), 169-197. [PubMed]
- Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*, 32(3), 193-254. [PubMed]
- Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79 (Pt 4), 471-491. [PubMed]
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics*, 33(2), 113-120. [PubMed]
- Weisstein, N., & Harris, C. S. (1974). Visual detection of line segments: An object-superiority effect. *Science*, 186(4165), 752-755. [PubMed]
- Wenger, M. J., & Ingvalson, E. M. (2002). A decisional component of holistic encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(5), 872-892. [PubMed]
- Wertheimer, M. (1923). Laws of organization in perceptual forms. First published as *Untersuchungen zur Lehre von der Gestalt II*, in *Psychologische Forschung*, 4, 301-350. Translation published in Ellis, W. (1938). *A source book of Gestalt psychology*. London: Routledge & Kegan Paul.
- Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology*, 1, 59-85.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81(1), 141-145.