Sensitivity to spatial frequency and orientation content is not specific to face perception

N. Rankin Williams [a,*], Verena Willenbockel [b], Isabel Gauthier [a]

[a] Department of Psychology, Vanderbilt University, Nashville, TN 37203, USA
[b] Department of Psychology, University of Victoria, Canada

ABSTRACT

Prior work using a matching task between images that were complementary in spatial frequency and orientation information suggested that the representation of faces, but not objects, retains low-level spatial frequency (SF) information [Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society of London, Series B Biological Sciences*, 352, 1203–1219]. In two experiments, we reexamine the claim that face perception is uniquely sensitive to changes in SF. In contrast to prior work, we used a design allowing the computation of sensitivity and response criterion for each category, and in one experiment, equalized low-level image properties across object categories. In both experiments, we find that observers are sensitive to SF and orientation changes for upright and inverted faces and non-face objects. Differential response biases across categories contributed to a larger sensitivity for faces, but even sensitivity showed a larger effect for faces, especially when faces were upright and in a front-facing view. However, when objects were inverted, or upright but shown in a three-quarter view, the matching of objects and faces was equally sensitive to SF changes. Accordingly, face perception does not appear to be uniquely affected by changes in spatial filter components.

Published by Elsevier Ltd.

1. Introduction

Broadly speaking, two categories of information are thought to be more critical for face than object perception: information about the configural relations between parts and the specific spatial frequency (SF) information present in images. Generally, studies report quantitative differences between face and object perception on measures designed to index how observers rely on these sources of information. For instance, a disadvantage for processing upside-down faces (a face inversion effect, see Rossion and Gauthier (2002) for review) has been used as an indirect measure of sensitivity to configural relations. But inversion typically also affects the perception of objects, just less so than it affects face perception (Rossion & Gauthier, 2002). Such evidence may not be strong enough to support the claim that face perception relies on one or several processes that are not available to object perception (McKone, Kanwisher, & Duchaine, 2007). Typically, such claims are made on the basis of qualitative differences between faces and non-face objects. In this work, we revisit prior claims that face perception differs qualitatively from that of objects in terms of its sensitivity to SF information (Biederman & Kalocsai, 1997).

There could be a process unique to face perception even if behavioral measures generally find only a *quantitative* difference between faces and objects. This would be the case if face perception also relies to some degree on part-based processes that are shared with generic object processing. Ideally, however, some tasks could be designed to be sensitive only to the process hypothesized to be face-specific, so that a *qualitative* behavioral difference can be documented. One measure that was suggested to reveal such a qualitative difference is the alignment effect in the composite task (Robbins & McKone, 2003; Young, Hellawell, & Hay, 1987). In this task, participants are asked to selectively attend to one part of a face made of the top and bottom halves of different faces, with these two halves aligned or misaligned. When the parts are aligned, participants have difficulty ignoring the irrelevant part of the composite[1]. However, a recent study showed that observers trained to individuate objects from a novel category also demonstrated an alignment effect in a composite task (Wong, Palmeri, & Gauthier, in press). While some hallmarks of face processing can be obtained only in expert observers, other effects once thought to be unique to faces have been obtained with objects in novice

---

[1] See Gauthier & Bukach (2007) and McKone & Robbins (2007) for a debate regarding different experimental designs to measure configural and holistic processing using composite stimuli.

* Corresponding author.
  E-mail address: Rankin.Williams@vanderbilt.edu (N.R. Williams).

observers. This is the case with the whole-part advantage: the finding that face parts studied in the context of a whole face are better recognized than in isolation (Tanaka & Farah, 1993). While the effect was originally obtained for faces and not houses, later studies reported a significant, albeit smaller, whole-part advantage in novice viewers with dogs, cars, and novel objects called Greebles (e.g., Tanaka & Gauthier, 1997).

The present study is an investigation of one of the rare behavioral effects so far only observed for faces. We call this effect the "Complementation Effect" (CE), and it indexes the sensitivity of face perception to manipulations of spatial filter components. Although this has been relatively less studied than other effects, face perception is reported to be highly sensitive to SF filtering (Fiser, Subramaniam, & Biederman, 2001; Goffaux, Gauthier, & Rossion, 2002) and to other types of manipulations of image format, such as contrast reversal (Gaspar, Bennett, & Sekuler, 2008; Hayes, 1988; Subramaniam & Biederman, 1997) and the use of line drawings (e.g., Bruce, Hanna, Dench, Healey, & Burton, 1992). These manipulations have a more limited impact on object recognition (Biederman, 1987; Biederman & Ju, 1988; Liu, Collin, Rainville, & Chaudhuri, 2000; Nederhouser, Yue, Mangini, & Biederman, 2007), suggesting that face and object perception may rely on different mechanisms and/or representations. Specifically, Biederman and Kalocsai (1997) explored the SF sensitivity of face perception. Complementary images were created by dividing the SF-by-orientation space into an 8×8 matrix and filtering out every odd diagonal of cells to form one version of an image and every even diagonal of cells to form the second image. These two versions of the same image are complementary in the sense that they do not overlap in any specific combination of SF and orientation (see Fig. 1). As might be expected, participants demonstrated a CE for faces, whereby they were poorer verifying and matching complementary faces relative to identical faces in both a name verification priming task and a same–different sequential matching task. But, perhaps more surprisingly, no CE was observed in either paradigm for common objects or chairs. Because the naming task was inherently confounded by task demands and the level of categorization (i.e., objects were named at the basic-level while famous faces were named at the subordinate-level), we are focusing here on understanding the face-object discrepancy observed via the sequential matching paradigm. Biederman and Kalocsai argue that this difference arises because the visual system represents faces and objects in distinct ways. They propose that non-face objects are stored as qualitative constructions of volumetric structural units (geons) that can be recovered from images based on non-accidental properties

found in an edge description of the object, devoid of the original SF image information (Biederman, 1987). In contrast, face representations are thought to preserve the specific information from V1-type cell outputs, accounting for why face perception is highly sensitive to SF–orientation manipulations.

Given that the CE was originally obtained for faces but not for non-face objects in novice observers (Biederman & Kalocsai, 1997), another study asked whether this effect may increase with perceptual expertise. Yue, Tjan, and Biederman (2006) trained participants with novel objects called blobs. All participants – those trained with blobs and those with no pre-testing exposure – showed robust CEs for faces and none for blobs. In addition, using fMRI these authors found that relative to an identical pair of images, a complementary pair of faces, but not blobs, reduced fMRI adaptation in the fusiform face area. The results of Biederman and Kalocsai (1997) and Yue et al. (2006) suggest that the CE is unique to faces. This is consistent with other work finding that the matching of objects such as chairs shows little sensitivity to manipulations of the overlap in SF content (Collin, Liu, Troje, McMullen, & Chaudhuri, 2004).

In the following experiments we revisit the question of whether the CE is unique to faces, guided by four main motivations. First, Biederman and Kalocsai (1997) and Yue et al. (2006) measured the CE by comparing accuracy in identical vs. complementary trials, when face or object identity was the same. The trials in which item identity (and thus the correct response) was different were pooled together, without being assigned to either condition (identical or complementary). Therefore, it is possible that observers applied different response criteria to face and non-face conditions tested in Biederman & Kalocsai and Yue et al.'s same–different matching tasks. Yue et al. reported errors collapsing over both same and different trials. While they reported no main effect for whether trials were same vs. different, it is nonetheless possible that an interaction with this factor approached significance and influenced the results.

Indeed, important differences in response biases between conditions, even when trials are not presented in different blocks, have been observed in other face processing studies and, when not accounted for, can lead to misleading conclusions (e.g., Cheung, Richler, Palmeri, & Gauthier, 2008; Wenger & Ingvalson, 2002). Therefore, to verify that the interaction between category and complementation is not due to differential response bias, we blocked trials by complementation condition so that two sets of different trials would be associated with identical vs. complementary
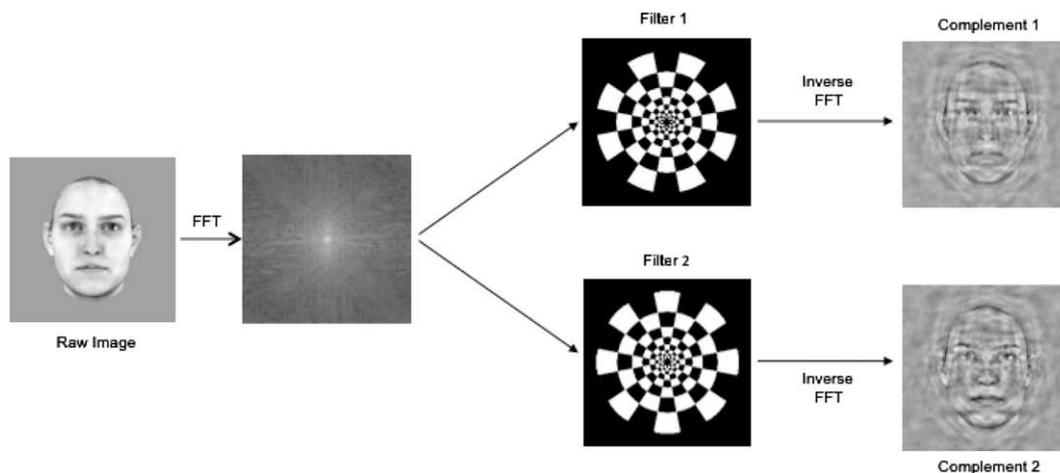
Fig. 1. Spatial frequency (SF) and orientation filtering. Two complementary images were created by filtering a single input image in the Fourier domain into an 8×8 radial matrix of SF–orientation information. Two separate filters were applied to preserve alternating combinations of the SF–orientation content of the original image. Thus, when returned to the spatial domain via inverse FFT, the complementary pair of images share no overlapping combinations of SF and orientation information.

conditions, allowing computation of discriminability and response bias. Second, in Experiment 1 we used an inversion manipulation to explore whether the CE can be attributed to configural processing, typically associated with face perception. While stimulus inversion is not a direct manipulation of configural processing, it is generally accepted that inversion affects the processing of configural information (Searcy & Bartlett, 1996; Tanaka & Sengco, 1997; Thompson, 1980). Third, prior comparisons of the CE between faces and objects made no attempt at matching the SF content of the original images across categories. It is possible that face perception is most sensitive to complementation because faces contain more information in a particular region of the SF space than control objects. Therefore, in Experiment 1 we match images of faces, cars and chairs in terms of low-level image properties, before applying the complementation filters. Finally, in Experiment 2 we investigate whether the symmetry of facial images used in prior work and in our Experiment 1 plays a role in the CE. It is possible that the radial symmetry of the filters is particularly disruptive to the encoding of symmetrical objects such as front facing images, because the filter would create symmetrical changes that may be especially likely to be interpreted as structural information (rather than alterations due to filtering).

2. Experiment 1

2.1. Methods

2.1.1. Participants

Thirty-four individuals (11 male, mean age 20 years) participated for a small honorarium or course credit. All participants had normal or corrected-to-normal visual acuity. The experiment was approved by the Institutional Review Board at Vanderbilt University, and all participants provided written informed consent.

2.1.2. Stimuli and material

Stimuli were digitized, eight-bit greyscale images of 15 faces with hair cropped (from the Max-Planck Institute for Biological Cybernetics in Tuebingen, Germany), 15 cars (all profile views obtained from www.tirerack.com), as well as 15 chairs (obtained from C. Collin, and used in Collin et al. (2004)). They were presented on a 21-in. CRT monitor (refresh rate = 100 Hz) using a Macintosh G4 computer running Matlab with the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997).

First, stimulus area was matched across object categories, such that the smallest rectangle containing each object occupied approximately 33% of the square window in which it was centered. The square window was presented in two sizes, with smaller sizes at 57 by 57 pixels and spanning approximately $1.98° \times 1.98°$ of visual angle, and larger sizes at 113 by 113 pixels and spanning approximately $3.95° \times 3.95°$ of visual angle.

Before filtering SF–orientation information, we equated a number of low-level image properties (i.e., luminance distributions and Fourier amplitude at each SF) using functions from the Spectrum, Histogram, and Intensity Normalization and Equalization (SHINE) program written with Matlab (see supplemental online information). First, the luminance histograms of the foregrounds and the backgrounds of all source images were collected, averaged separately across the set, and applied to each stimulus. Second, we obtained the average Fourier amplitude spectrum across the set and equated the rotational average amplitude for each SF across all stimuli. The equalization steps were performed iteratively 25 times to reach a high degree of simultaneous equalization of luminance histograms and Fourier amplitudes.

After images were SHINEd, they were filtered with a method identical to that used by Biederman and Kalocsai (1997) and Yue et al. (2006). The original images were Fourier transformed and filtered using two complementary filters (Fig. 1). Each filter eliminated the highest (above 181 cycles/image) and lowest (below 12 cycles/image, corresponding to approximately 7.5 cycles per face width (c/fw)) SFs. Note that prior work suggests that face recognition relies primarily on a middle band of SFs (approximately 8–16 c/fw – see Collin et al. (2004) for recent review). The surviving area of the Fourier domain was divided into an 8×8 matrix with eight orientations (increasing in successive steps of 22.5°) by eight SFs (covering four octaves in steps of 0.5 octaves). This manipulation created two complementary pairs of images, whereby every other 32 SF–orientation combinations in a radial checkerboard pattern in the Fourier domain was ascribed to one image and the remaining combinations were assigned to the complementary member of that image pair. As such, both complementary members of a pair contained all eight SFs and all eight orientations but in unique combinations. Thus, the two complementary images shared no common information about the objects in the Fourier domain. After images were filtered in the Fourier domain, they were converted back to images in the spatial domain via the inverse FFT. The final processed images (see Fig. 2) were presented at a viewing distance of 58 cm.

2.1.3. Procedure

We used a $3 \times 2 \times 2$ within-subjects design, with factors being Category (face, car, or chair identity), Complementation (identical or complementary) and Orientation (upright or inverted). Trials were blocked according to SF composition (identical or complementary) in order to conduct signal detection analysis, and the visual angle of the image always varied from study to probe, either 2° or 4 ° of visual angle. There were six blocks (three identical, three complementary) of 192 trials (64 face trials, 64 car trials and 64 chair trials), where image Category and Orientation varied randomly within a block. Block order was randomized across subjects. Participants were given 12 practice trials and offered a break every 64 trials.

On each trial, participants judged whether a pair of sequentially presented images (either two faces, two cars, or two chairs) was of the same identity. Relative to the study image, the probe image could be (a) the same identity and the same SF (i.e., the exact image), (b) the same identity and a complementary SF, or (c) a different exemplar altogether. The manipulation of interest was therefore SF overlap or complementarity – as opposed to the SF content *per se* – of two sequentially flashed images. Participants were instructed to make their judgments based on identity alone, regardless of differences in image size or SF content (described to subjects as "blurriness"). As in Yue et al. (2006), image size always differed from study to probe (2–4 deg or 4–2 deg), so that no part of the effect could be attributed to image matching.

Each trial began with a 500 ms fixation cross, followed by a target stimulus (face, car or chair) in the center of the screen for 200 ms. After a 300 ms inter-stimulus-interval a probe stimulus of the same category appeared for 200 ms. Participants had to make a same/different judgment regarding the identity of the target and probe within 1800 ms.

Following the matching task with filtered images, all participants completed a test of car expertise to quantify their skill at car identification (Curby, Glazek, & Gauthier, 2009; Gauthier, Curby, & Epstein, 2005; Gauthier et al., 2000; Grill-Spector, Knouf, & Kanwisher, 2004; Rossion, Collins, Goffaux, & Curran, 2007; Xu, 2005). We did not explicitly recruit car experts, but the data were acquired to explore whether natural variation in car expertise may account for a potential difference between cars and chairs. In this task, participants made same/different judgments on car images (at the level of make and model, regardless of year) and bird images (at the level of species). For each of 112 car trials and 112
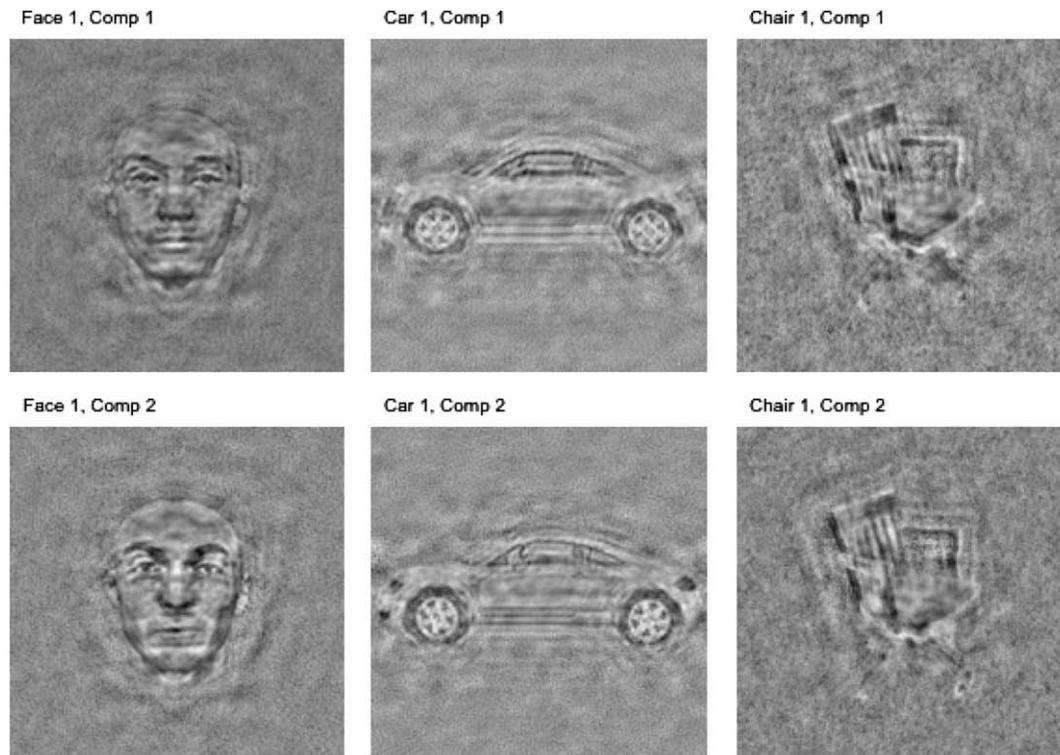
Fig. 2. Example images from Experiment 1. Displayed are pairs of ~~filtered and SHINEd~~ exemplars from each object category (face, car, and chair).

bird trials, the first stimulus appeared for 1000 ms, followed by a 500 ms mask. A second stimulus then appeared and remained visible until a same/different response was made or 5000 ms elapsed. A separate sensitivity score was calculated for cars (Car d') and birds (Bird d'). The difference between these measures (Car d'–Bird d') yields a Car Expertise Index for each participant. Participants span a limited range of car performance, comparable to the range we observed for birds (Car d' range = 0.31–2.25; Bird d' range = 0.55–1.55; Car Expertise Index range = −0.52 to 1.42). Expertise data were not obtained for two individuals who failed to take the Expertise test due to personal time constraints. We made no effort to recruit participants that were experts with cars, so the range of car expertise was limited relative to prior work focusing on car expertise.

2.2. Results

Fig. 3a presents the mean results in each condition for three dependent variables. First, we consider accuracy on same trials only to provide an index that is comparable to what was used in prior studies (Biederman & Kalocsai, 1997; Yue et al., 2006). Second, we look at discriminability (d') and response criterion (C), incorporating both the same and the different trials into our design. Paired t-tests revealed a significant CE in every condition using either accuracy, d', or response criterion (all $ps < 0.0001$).

Data were analyzed using a $3 \times 2 \times 2$ ANOVA with within-subject factors being Category (face, car, or chair), Complementation (identical or complementary), and Orientation (upright or inverted).

Considering first accuracy on same trials, all main effects and interactions were significant (all $ps < 0.0001$). Performance was better for identical pair trials than complementary pair trials ($F_{(2, 33)} = 203.28$), and for inverted than upright trials ($F_{(2, 33)} = 52.80$). When investigating the main effect of Category type ($F_{(2, 33)} = 31.46$) via Bonferroni post hoc tests (per-comparison alpha level = 0.017), we found that performance was better for chairs and cars relative to faces, and equivalent for chairs and cars.

To investigate the three-way interaction between Category, Complementation and Orientation ($F_{(2, 66)} = 36,67$), we conducted a two-way ANOVA directly on participants' mean complementation scores for each condition (Identical–Complementary, see Fig. 3b). Again, both main effects (those of Category and Orientation) and the interaction were significant (all $ps \leqslant 0.0001$). Of particular interest is the Category × Orientation interaction ($F_{(2, 66)} = 36.67$), which we followed up with post hoc tests (Bonferroni, using a per-comparison alpha = 0.0056). Only faces saw their CE reduced by inversion. In addition, upright faces showed a larger CE than both upright cars and chairs (with no difference between cars and chairs themselves), and this pattern persisted for inverted stimuli. Note, however, that there was a significant CE in all conditions (all $ps \leqslant 0.0001$).

Accuracy for different trials is reported in Appendix A, along with performance rates by collapsing same- and different-identity trials together as in Yue et al. (2006). Collapsing over all trials, the complementation effect remains significant for all conditions, as revealed through a significant three-way interaction (Category × Orientation × Complementation: $F_{(2, 33)} = 13.00$, $p < 0.0001$), and confirmed via Bonferroni post hoc analyses. Note however that this analysis is problematic given the large differences in bias across conditions and it is only included to compare with prior findings.

How do these results compare with a signal detection analysis that uses both same and different trials to separate discriminability from response criterion? When analyzing d' in a $3 \times 2 \times 2$ ANOVA, again, all main effects and interactions were significant (all $ps < 0.01$). We emphasize here only those effects that depart qualitatively from the analysis on accuracy. When looking at the main effect of Category on sensitivity ($F_{(2, 33)} = 29.54$, $p < 0.0001$) via a Bonferroni post hoc comparison test (per-comparison alpha level = . 017), we found that performance was better for chairs relative to faces and cars, and equivalent for face and car trials.

To investigate the three-way interaction between Category, Complementation, and Orientation ($F_{(2, 33)} = 4.58$), we conducted a two-way ANOVA directly on participants' mean complementation
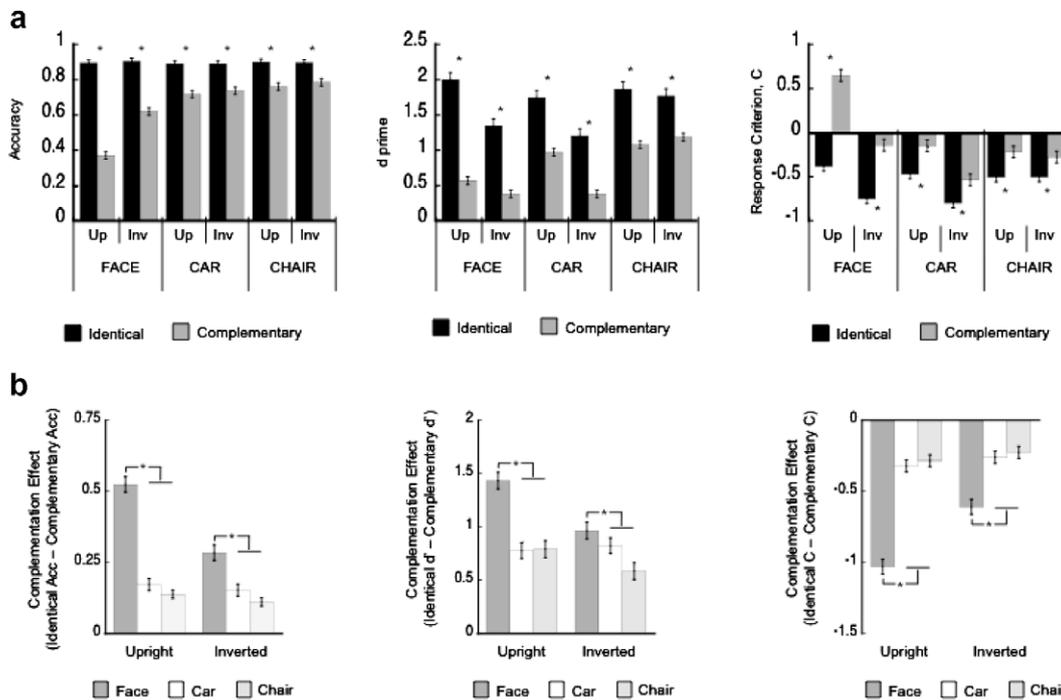
Fig. 3. Experiment 1 results (*N* = 34). (a) Mean values for same–different matching of identical and complementary faces, cars, and chairs presented either upright or inverted are represented using three dependent measures: accuracy for same trials, sensitivity (*d′*), and response criterion (*C*). For each orientation of all stimulus categories using both accuracy and *d′* measures, subjects were significantly better matching identical relative to complementary pairs of images. Error bars represent the standard error of the mean. (b) The CE (Identical–Complementary) associated with accuracy, *d′* and *C* measurements is plotted for upright and inverted faces, cars, and chairs. Error bars represent standard error of the mean.

scores for each condition (Identical–Complementary, see Fig. 3b). Again, both main effects and interaction were significant (Category (*p* < 0.0001), orientation (*p* < 0.01), Category × Orientation (*p* = 0.01)). Of particular interest is the Category × Orientation interaction (*F*(2, 66) = 4.58). Post hoc tests (Bonferroni, alpha = 0.0056) showed only faces saw their CE reduced by inversion. In addition, while for upright stimuli faces showed a larger CE than both cars and chairs (no difference between non-face categories), for inverted stimuli the face CE was not significantly larger than that for cars or chairs.

It is easy to appreciate from Fig. 3 that response criterion contributed to exaggerate the CE in accuracy for upright faces. A 3 × 2 × 2 ANOVA shows all main effects and interactions to be significant (all *p*s < 0.0001). To investigate the three-way interaction (*F*(2, 66) = 13.245) we performed a 2 × 2 ANOVA directly on CEs (Identical–Complementary). All main effects and interactions were significant (all *p*s < 0.0001). Of interest is the interaction between Category and Orientation (*F*(2, 66) = 13.245), which we followed up with post hoc tests (Bonferroni, alpha = 0.0056). It revealed that only for faces was the CE in response criterion influenced by inversion. In addition, for both upright and inverted stimuli, there was a larger CE for faces relative to either cars or chairs (which were statistically equivalent).

For brevity, analyses on response times are not reported here but the mean values for all conditions for same-identity trials, different-identity trials, and same/different combined trials are shown in Appendix A. They were generally consistent with those for *d*-prime and did not suggest any speed–accuracy tradeoffs.

2.3. Discussion

In Experiment 1, we observed a CE in all conditions (faces, cars and chairs, both upright and inverted), even when analyzing only accuracy on same trials. As such, with stimuli equated in a number of low-level properties (luminance histograms and Fourier amplitudes), we failed to replicate prior results in which no CE was observed for objects (Biederman & Kalocsai, 1997; Yue et al., 2006). We did, however, observe as in these studies that the CE for faces was larger than that for objects. Importantly, this difference was exaggerated by the use of accuracy for same trials as a dependant variable, as revealed through an important difference in participants' response criterion between faces and objects. While observers were generally biased to respond "same" in most conditions of this experiment, they showed a unique bias to say "different" for complementary images of upright faces. This bias is independent of the ability of observers to judge whether the complementary images are the same or not. Indeed, the same response bias is not observed in other conditions where performance is either better (e.g., complementary chairs) or poorer (e.g., complementary inverted cars or faces) than for upright faces.

Given this important effect of response bias, it is more appropriate to compare the CE for faces and objects using sensitivity (*d′*). Using *d′*, we still find the largest CE for upright faces, compared not only to other object categories but also to inverted faces, suggesting that the larger CE for faces is not fully accounted for by low-level differences that may persist between faces and objects despite the application of SHINE. That is, while we have eliminated many of these differences across categories using SHINE, it remains possible that this procedure affected information that was more diagnostic for the discrimination of faces than other objects. The inversion effect suggests that the magnitude of the face CE is influenced by configural processing, well known to be more available for upright than inverted faces (Searcy & Bartlett, 1996; Tanaka & Sengco, 1997; Thompson, 1980). When only inverted stimuli are considered, the face CE is no longer larger than the car CE. Note that cars showed an overall inversion effect comparable to that found with faces (when only cars and faces are included in an ANOVA using *d′*, the Category × Orientation effect is not significant,

$p = 0.11$). Cars are highly familiar and mono-oriented objects, and car inversion effects with car novices have been observed in other tasks (Gauthier et al., 2000). Despite this, inversion for cars did not interact with the magnitude of the CE.

3. Experiment 2

In Experiment 2, we address the following issues. First, because we found a CE for objects even using accuracy as a dependent measure, which was not the case in Biederman and Kalocsai (1997) and Yue et al. (2006), we wanted to ensure that the SHINE procedure was not the cause of the CE for objects. We therefore filtered images of faces and chairs (the category that was maximally different from faces in Experiment 1) without using SHINE, as in the previous studies. Second, while prior work used actual face photographs, Experiment 1 employed images from a face database wherein each face represents a linear combination, or morphing, of a set of prototypical faces created from cylindrical 3-D laser scanning (Blanz & Vetter, 1999; Troje & Bülthoff, 1996). To test for the generalization of our effects, we therefore used a different set of faces: regular face photographs from the CVL database. Finally, we wanted to investigate the possible role of image symmetry in the CE. In Experiment 1, faces were shown in a symmetrical front-facing view, whereas cars were displayed in profile and chairs rotated towards the right about 45°. It is possible that these differences contributed to the larger CE for upright faces, because the radially symmetrical SF–orientation filter applied would have produced symmetrical changes on the faces that may have been especially likely to be interpreted as changes in the shape of the faces. Therefore, in Experiment 2 we measured the CE for 0° and 45° views of both faces and chairs.

3.1. Methods

3.1.1. Participants

Thirty individuals (15 male, mean age 19 years) participated for a small honorarium or course credit. All participants had normal or corrected-to-normal visual acuity. The experiment was approved by the Institutional Review Board at Vanderbilt University, and all participants provided written informed consent.

3.1.2. Stimuli and material

Stimuli were images (30 faces with hair cropped from the Computer Vision Laboratory at the University of Ljubljana in Ljubljana, Slovenia, and 30 chairs obtained from C. Collin and used in Collin et al. (2004)). For each object category, half of the images were front-on image views (0° rotation) while the other half were quarter views (45° rotation – see Fig. 4). All images were filtered into complementary pairs as in Experiment 1, and then transformed to two different sizes, either 113 by 113 pixels ($\sim$3.95° × 3.95° of visual angle) or 226 × 226 pixels ($\sim$7.85° × 7.85° of visual angle). All images were presented at a viewing distance of 58 cm.

3.1.3. Procedure

This experiment employed a 2 × 2 × 2 repeated-measures design with the following factors: Category (face or chair), ~~SF orientation content~~ (identical or complementary), and Viewpoint (0° view or 45° view). The viewpoint of the stimuli varied randomly across trials, though both stimuli within a trial were always shown in the same view. As before, image size differed from study to probe to reduce image matching, and trials were blocked according to SF content (identical or complementary) to allow for the computation of sensitivity and response bias.

Participants completed eight blocks of 128 trials each: four blocks of identical SF–orientation pairs and four blocks of complementary SF–orientation pairs. Stimulus category and viewpoint varied randomly within a block, allowing 256 trials for each combination (i.e., 0° faces, 0° chairs, 45° faces, 45° chairs). Block order was randomized across subjects. Each subject began with 15 practice trials, and breaks were offered every 64 trials. Participants determined whether the probe stimulus was the same or different identity from the target stimulus.
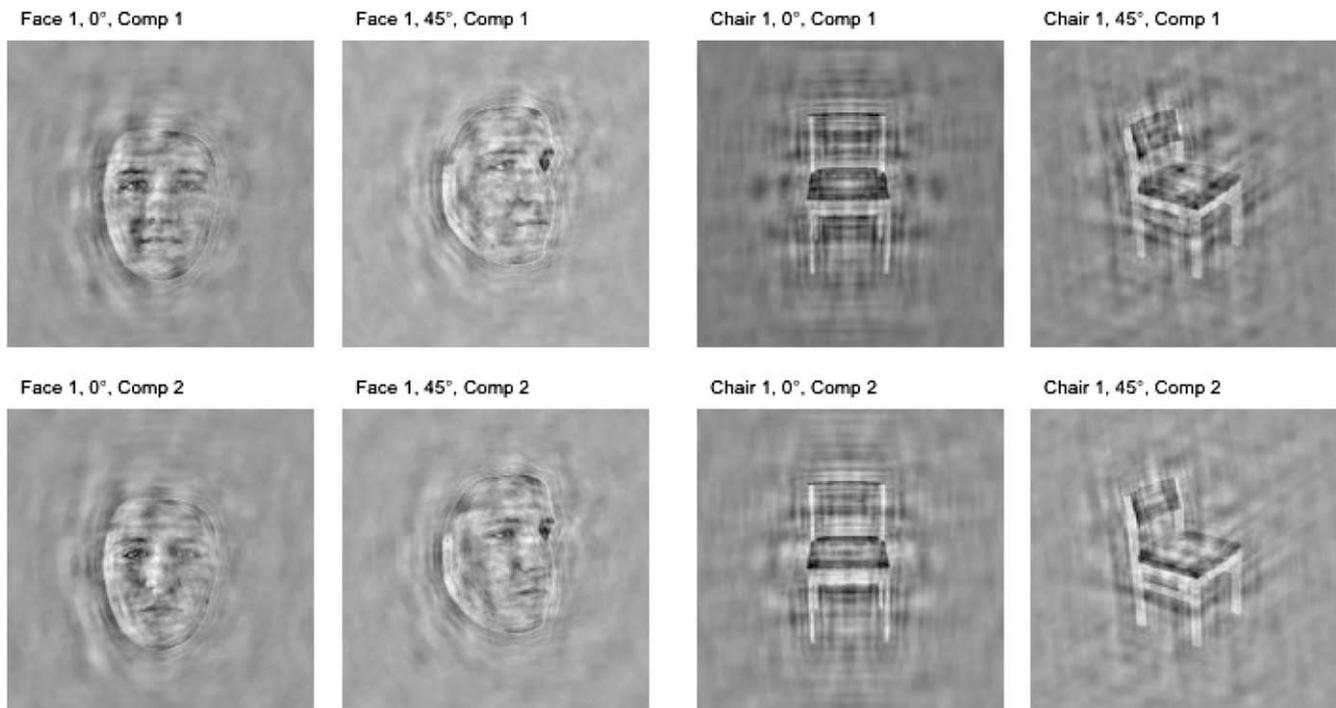
Fig. 4. Example images from Experiment 2. Displayed are pairs of filtered images from each target category (0°- and 45°-view faces and chairs).

3.2. Results

Fig. 5 presents accuracy on same trials, discriminability (d') and response criterion (C). Paired t-tests revealed a significant CE in every condition using accuracy, d', and C measures (all $ps < 0.02$). For each dependant variable, the results were further analyzed using a $2 \times 2 \times 2$ ANOVA with factors (all within-subjects) being Category (face, chair), SF orientation manipulation (identical, complementary), and Viewpoint ($0°$, $45°$).

For Accuracy with same-identity trials, all main effects and interactions were significant (all $ps < 0.0002$). Performance was better overall for chairs than faces ($F_{acc}(1, 29) = 32.14$), for $45°$ than $0°$ views ($F_{acc}(1, 29) = 53.17$), and for identical than complementary pairs ($F_{acc}(1, 29) = 182.74$).

To investigate the three-way interaction ($F_{acc}(1, 29) = 33.94$), we conducted 2×2 ANOVAs directly on CE scores (Identical–Complementary, see Fig. 5b). All main effects and interaction were significant ($p < 0.0001$). We were especially interested in the Category $\times$ Viewpoint interaction ($F_{acc}(2, 29) = 33.94$), which we followed with Bonferroni post hoc tests (per-comparison alpha = 0.0125). The CE for faces was greater than for chairs regardless of viewpoint. In addition, while face stimuli showed a reduced CE when rotated $45°$, chair stimuli showed comparable CEs for $45°$ views than for $0°$ views.

The Appendix A gives accuracy for different trials, as well as performance with identical and complementary filters collapsing over same- and different-identity trials. As in Experiment 1, there is a significant three-way interaction (Category $\times$ Orientation $\times$ Complementation: $F(2, 29) = 6.13$, $p = 0.0194$). Bonferroni post hoc analyses (per-comparison alpha = 0.0125) reveal a significant CE for faces regardless of viewpoint, while the CE was not significant for chairs. Note that given the large response bias differences across categories, this analysis is only included as a reference.

When analyzing d' in a $2 \times 2 \times 2$ ANOVA, the main effects of Category ($Fd'(1, 29) = 234.22$, $p < 0.0001$) and Complementation ($Fd'(1, 29) = 55.12$, $p < 0.0001$) remained significant. Viewpoint showed only a marginally significant effect ($Fd'(1, 29) = 3.68$, $p = 0.06$). The only other significant effect was an interaction between Category and Complementation ($Fd'(1, 29) = 5.41$, $p = .03$). Post hoc tests (Bonferroni, per-comparison alpha = 0.0125) revealed that with both sets of stimuli (identical and complementary) subjects were better with chairs than faces, although there was a significant CE for both categories.

To explore the effect of image symmetry using d', we also conducted a two-way ANOVA (Category $\times$ Viewpoint) on the CE for this measure. Neither the effect of viewpoint ($Fd'(1, 29) < 1$, n.s.) nor the interaction effect ($Fd'(1, 29) < 1$, n.s.) were significant using d'. It is interesting to note, therefore, that viewpoint has less of an impact with d' than with accuracy.

Considering response criterion (C) in a $2 \times 2 \times 2$ ANOVA, all main effects and interactions were significant ($ps < 0.001$) except the main effect of category ($F_C(1, 29) = 2.80$, $p = 0.10$). As in Experiment 1, participants demonstrated a particularly large bias to say "different" for complementary pairs of $0°$ view face images. To investigate the significant three-way ANOVA, we computed a 2×2 ANOVA directly on CE (Identical–Complementary). The interaction between Category and Viewpoint ($F_C(1, 29) = 12.96$, $p = 0.001$) reflected the fact that there was, according to post hoc tests (Bonferroni, per-comparison alpha = 0.0125), a significant effect of viewpoint for faces but not for chairs (as is easily appreciated from Fig. 5b). Response times for Experiment 2 are reported in Appendix A.
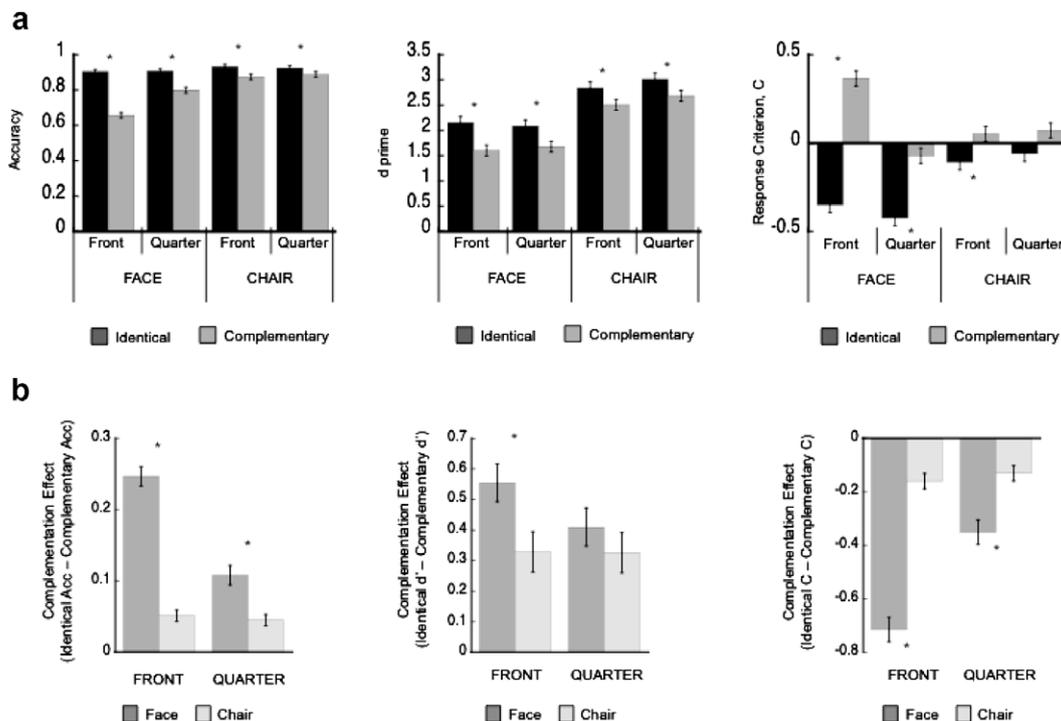
Fig. 5. Experiment 2 results ($N = 30$). (a) Mean values for same–different matching of identical and complementary faces and chairs at either $0°$ rotation or $45°$ rotation are represented using three dependent measures: accuracy for same trials, sensitivity (d'), and response criterion (C). Error bars represent standard error of the mean. (b) The CE (Identical–Complementary) associated with accuracy, d' and C measurements is plotted for $0°$ and $45°$ views of faces and chairs. Error bars represent standard error of the mean.

3.3. Discussion

In summary, Experiment 2 is consistent with Experiment 1, replicating and extending several findings. We find CEs for both faces and objects, and differences between faces and objects are amplified when using accu— by the presence of a differential response bias. Moreover, viewpoint reduced the response bias difference between identical and complementary trials for faces but not for chairs.

4. General discussion

Using the same SF–orientation filtering mechanism as employed here, Biederman and Kalocsai (1997) concluded that face recognition, but not object recognition, is sensitive to changes in the spatial filter components of the image. On this basis they suggested that faces and objects are stored in qualitatively unique ways in the brain: faces are represented as a collection of outputs of V1-type cells preserving their original SF- and orientation-selective information, while objects are represented as structural units composed of lines and edges derived from harmonics that transcend scale and are devoid of SF information.

In contrast, our results indicate that significant CEs can also be obtained with chairs and cars (regardless of whether they are processed to match the low-level properties of faces). The CE is obtained with upright and inverted objects, regardless of viewpoint. We suggest, therefore, that the CE is not unique to faces.

Why was the CE not obtained with objects in prior work? While it is always difficult to interpret null effects, we should note that our results are generally consistent with those of Biederman and Kalocsai (1997) and Yue et al. (2006) in that we found the CE to be larger for faces than objects. It is possible that that ignoring potential differences in response bias across conditions may have dampened the CE for objects, as in our Experiment 2 but we note that in Experiment 1 we found CEs for non-face objects regardless of analysis method. Importantly, our findings suggest that the magnitude of the CE is influenced by factors that are unlikely caused by qualitative difference in underlying representations across categories. First, there is an important difference in response bias between faces and objects, which inflates the size of the CE for faces. We do not have a theoretical account of this difference, but we note that such a differential bias between faces and objects has been observed in other tasks (Cheung et al., 2008; Wenger & Ingvalson, 2002). An anonymous reviewer of this paper suggested that the differential bias effect for faces may be based on our blocking by complementarity, so that we could unambiguously attribute different trials to the identical or the complementary condition, and that a similar bias difference may not have been found in Biederman and Kalocsai (1997) and Yue et al. (2006) who only blocked trials by category. To test this account we re-analyzed data from an experiment in a different paper (Williams & Gauthier, submitted for publication). In that study ($N = 39$) which used the same filters as the present work, there were upright and inverted face and car trials blocked by category and randomized by complementarity as in prior work. To assign the "different" trials to complementary or identical conditions, we simply followed the same rule as for the "same" trials: namely, whether the *filter* was identical or complementary. This re-analysis replicated the large difference in response bias between faces and cars even when trials were randomized as in prior work (Mean C for upright identical faces = -0.188; upright complementary faces = 0.740; upright identical cars = -0.179; upright complementary cars = 0.221; interaction between category and complementarity: $F(1, 38) = 155.97$, $p < 0.001$). This allows us to reject the possibility that the differential bias effects for upright faces obtained in our experiments can be explained by blocking by complementarity. It remains possible that the bias effect has a perceptual basis: in fact, it may seem unlikely that a response bias would be relatively specific to front-view upright faces. Nonetheless, experience with a category could also influence decisional aspects of these matching judgments. This could be tested in future work in a framework that could explicitly address the nature of the interaction between the task relevant identity and the task irrelevant filtering, such as the General Recognition Theory (GRT; Ashby & Townsend, 1986).

In addition, another factor that could help explain the discrepancy with prior work is that in the present study, we found that both inversion and rotation in depth reduced the CE for faces. The objects manipulated in prior studies may not have been shown in a symmetrical view as were the faces (e.g., the blobs used in Yue et al., 2006). Therefore, differences in viewpoint or symmetry between faces and objects, coupled with the presence of response bias, may have contributed to reducing the CE for objects in prior work. Clearly other differences could have been influential, as well; for example the CEs we observed in Experiment 2 were considerable smaller than those from Experiment 1, an effect not predicted but which could result from any of several confounded differences in methods (use of SHINE, different stimulus sets, different image sizes).

Nonetheless, beyond the general cost of complementation, our results are consistent with the idea that the matching of front views of faces is more affected by manipulations of the spatial information than is object matching. In contrast to Collin et al. (2004) who found no inversion effect on the sensitivity to SF overlap for faces, we found inversion to reduce the CE for faces but not that for cars or chairs in both our experiments. This suggests that the magnitude of the CE may depend on configural processing. Indeed, while inversion does not *only* affect configural processing, an effect of inversion limited to face perception is more suggestive that configural processing is involved. In recent work, inversion effects observed for faces and not for objects have been replicated in expert observers (Curby et al., 2009), and configural processing has also been found to increase with expertise (Gauthier & Tarr, 2002; Wong et al., in press). Therefore, an inversion effect for faces and not for objects is consistent, albeit indirectly, with a role of expertise in the CE. We found that rotation in depth also reduced the CE for faces, and this manipulation has been found not to interact with configural processing (McKone, 2008). The CE could be affected by inversion for the same core reason as it is reduced by rotation in depth, or the two effects could have unique underlying causes. Future work should address these questions and, in particular, more intentionally test the role of expertise in the CE, because expertise effects should be particularly reduced for very unfamiliar inverted views.

In summary, we found that the CE can be obtained with non-face objects. Using a measure of discriminability to reduce the role of important responses biases, we found that the CE was larger in magnitude for faces than for objects, and most pronounced for upright front views of faces. With inversion, or rotation in depth, the effect of complementation was no longer face-specific.

Appendix A

Accuracy and response times (ms; correct trials only) for "same"- and "different"-identity trials, as well as for identical and complementary filters collapsed across same- and different-identity trials, in Experiments 1 and 2.

Experiment 1: accuracy.

	Face		Car		Chair	
	Upright	Inverted	Upright	Inverted	Upright	Inverted
Identical, same trials	0.895	0.904	0.891	0.891	0.900	0.897
Identical, different trials	0.713	0.469	0.633	0.426	0.652	0.641
Identical, all trials	0.804	0.687	0.762	0.658	0.776	0.769
Complementary, same trials	0.371	0.620	0.719	0.738	0.762	0.786
Complementary, different trials	0.788	0.521	0.611	0.384	0.619	0.615
Complementary, all trials	0.580	0.571	0.665	0.561	0.691	0.701

Experiment 1: response time (ms) for correct trials.

	Face		Car		Chair	
	Upright	Inverted	Upright	Inverted	Upright	Inverted
Identical, same trials	468	504	485	505	464	492
Identical, different trials	567	582	582	596	609	583
Identical, all trials	518	543	533	550	537	538
Complementary, same trials	578	600	576	589	559	592
Complementary, different trials	549	594	599	613	603	644
Complementary, all trials	564	597	588	601	581	618

Experiment 2: accuracy.

	Face		Chair	
	0° View	45° View	0° View	45° View
Identical, same trials	0.904	0.907	0.932	0.926
Identical, different trials	0.750	0.717	0.888	0.904
Identical, all trials	0.827	0.812	0.910	0.915
Complementary, same trials	0.657	0.799	0.874	0.887
Complementary, different trials	0.851	0.759	0.889	0.903
Complementary, all trials	0.754	0.779	0.882	0.895

Experiment 2: response time (ms) for correct trials.

	Face		Chair	
	0° View	45° View	0° View	45° View
Identical, same trials	523	492	480	466
Identical, different trials	551	597	525	545
Identical, all trials	537	544	502	506
Complementary, same trials	617	552	512	510
Complementary, different trials	602	612	548	535
Complementary, all trials	609	582	530	522

References

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. Psychological Review, 93, 154–179.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115–147.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. Cognitive Psychology, 20, 38–64.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. Philosophical Transactions of the Royal Society of London, Series B Biological Sciences, 352, 1203–1219.

Blanz, V., Vetter, T. (1999). A morphable model for the synthesis of 3D faces. In Presented at Proceedings of the 26th annual conference on Computer graphics and interactive techniques.

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 1, 443–446.

Bruce, V., Hanna, E., Dench, N., Healey, P., & Burton, M. (1992). The importance of "mass" in line drawings of faces. Applied Cognitive Psychology, 6, 619–628.

Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. Journal of Experimental Psychology: Human Perception and Performance, 34, 1327–1336.

Collin, C. A., Liu, C. H., Troje, N. F., McMullen, P. A., & Chaudhuri, A. (2004). Face recognition is affected by similarity in spatial frequency range to a greater degree than within-category object recognition. Journal of Experimental Psychology: Human Perception and Performance, 30, 975–987.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. Journal of Experimental Psychology: Human Perception and Performance, 35, 94–107.

Fiser, J., Subramaniam, S., & Biederman, I. (2001). Size tuning in the absence of spatial frequency tuning in object recognition. Vision Research, 41, 1931–1950.

Gaspar, C. M., Bennett, P. J., & Sekuler, A. B. (2008). The effects of face inversion and contrast-reversal on efficiency and internal noise. Vision Research, 48, 1084–1095.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? Cognition, 103, 322–330.

Gauthier, I., Curby, K. M., & Epstein, R. (2005). Activity of spatial frequency channels in the fusiform face-selective area relates to expertise in car recognition. Cognitive and Affective Behavioral Neuroscience, 5, 222–234.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. Journal of Experimental Psychology: Human Perception and Performance, 28, 431–446.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. Journal of Cognitive Neuroscience, 12, 495–504.

Goffaux, V., Gauthier, I., & Rossion, B. (2002). Spatial scale contribution to early visual differences between face and object processing. Cognitive Brain Research, 16, 416–424.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within category identification. *Nature Neuroscience, 7*, 555–562.

Hayes, A. (1988). Identification of two-tone images; some implications for high- and low-spatial-frequency processes in human vision. *Perception, 17*, 429–436.

Liu, C. H., Collin, C. A., Rainville, S. J. M., & Chaudhuri, A. (2000). The effects of spatial frequency overlap on face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 729–743.

McKone, E. (2008). Configural processing and face viewpoint. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 310–327.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*, 8–15.

McKone, E., & Robbins, R. (2007). The evidence rejects the expertise hypothesis: reply to Gauthier & Bukach. *Cognition, 103*, 331–336.

Nederhouser, M., Yue, X., Mangini, M. C., & Biederman, I. (2007). The deleterious effect of contrast reversal on recognition is unique to faces, not objects. *Vision Research, 47*, 2134–2142.

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

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

Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization processes. *Journal of Cognitive Neuroscience, 19*, 543–555.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews. Developmental Psychobiology, 1*, 63–75.

Searcy, J. M., & Bartlett, J. C. (1996). Inversion and processing of component and spatial–relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 904–914.

Subramaniam, S., & Biederman, I. (1997). Does contrast reversal affect object identification? *Investigative Ophthalmology and Visual Science, 38*, 998.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal Experimental Psychology, 46*, 225–245.

Tanaka, J. W., & Gauthier, I. (1997). Expertise in object and face recognition. In Goldstone, Medin, & Schyns (Eds.). *Psychology of Learning and Motivation Series, Special Volume: Perceptual Mechanisms Learning* (Vol. 36, pp. 83–125). San Diego: Academic Press.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*, 583–592.

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

Troje, N., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research, 36*, 1761–1771.

Wenger, M. J., & Ingvalson, E. M. (2002). A decisional component of holistic encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 872–892.

Williams, N. R., Gauthier, I. (submitted for publication). Perceptual expertise with objects predicts hallmark of face perception.

Wong, A. C.-N., Palmeri, T. J., Gauthier, I. (in press). Conditions for face-like expertise with objects: Becoming a Ziggerin expert – but which type? *Journal of Psychological Sciences.*

Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex, 15*, 1234–1242.

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

Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special? *Vision Research, 46*, 3802–3811.