

Stimulus Type, Level of Categorization, and Spatial-Frequencies Utilization: Implications for Perceptual Categorization Hierarchies

Assaf Harel and Shlomo Bentin
Hebrew University of Jerusalem

The type of visual information needed for categorizing faces and nonface objects was investigated by manipulating spatial frequency scales available in the image during a category verification task addressing basic and subordinate levels. Spatial filtering had opposite effects on faces and airplanes that were modulated by categorization level. The absence of low frequencies impaired the categorization of faces similarly at both levels, whereas the absence of high frequencies was inconsequential throughout. In contrast, basic-level categorization of airplanes was equally impaired by the absence of either low or high frequencies, whereas at the subordinate level, the absence of high frequencies had more deleterious effects. These data suggest that categorization of faces either at the basic level or by race is based primarily on their global shape but also on the configuration of details. By contrast, basic-level categorization of objects is based on their global shape, whereas category-specific diagnostic details determine the information needed for their subordinate categorization. The authors conclude that the entry point in visual recognition is flexible and determined conjointly by the stimulus category and the level of categorization, which reflects the observer's recognition goal.

Keywords: perceptual categorization, levels of categorization, spatial frequencies, face processing, object recognition

Supplemental materials: <http://dx.doi.org/10.1037/a0013621.supp>

An evident difference between visual processing of objects and faces is the level at which they are categorized by default, that is, how they are perceived without task constraints. Whereas in most cases, objects are perceptually categorized faster at the basic level (e.g., a car) than at the subordinate level (e.g., a 1965 Ford Mustang; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), faces are seemingly categorized equally fast at basic (a face) and subordinate levels (George Bush; e.g., Tanaka, 2001).¹ This difference might reflect the different functionality and ecological importance of subordinate distinctions within the face and other object categories, as well as the different amount of knowledge and experience that people have in recognizing exemplars within each category. In line with the latter account, expert knowledge with within-category diagnostic features might overcome the added perceptual difficulty that characterizes subordinate-relative to basic-level categorization (Hamm & McMullen, 1998; Jolicoeur, Gluck, & Kosslyn, 1984) and change the entry point of the percept from basic to subordinate (subordinate shift; Tanaka & Taylor, 1991). In fact, it stands to reason that most behavioral and neural signatures that distinguish the categorization of faces and objects of different kinds (biological categories as well as artifacts) reflect differences in processes supporting within-category dis-

tinctions (Gauthier, 2000; Tarr & Gauthier, 2000). Therefore, exploring factors that modulate these category-selective signatures could provide insights about both the specific computations and the relevant perceptual information that underlie face and object recognition.

The concept of a basic level in perceptual categorization implies both a maximization of perceptual differences between categories and a minimization of the differences within a category (Rosch et al., 1976). Exemplars of a well-defined category usually share a canonical structure² comprising a typical shape (most of the time referring to contour), as well as a typical organization of parts. For well-defined basic-level categories, this structure is well distinguished from the structure shared by stimuli forming a different category. Therefore, basic-level categorization is most of the time based on the stimulus's shape that highlights the typicality of the exemplar (Rosch et al., 1976) and the shared organization of the exemplar's parts, which, in conjunction with the shape, contain the perceptual typicality of the category (Tversky & Hemenway, 1984). Airplanes, for example, share a typical shape that is very different from the typical shape shared, for example, by cars. Airplanes also share a spatial configuration of parts that is different from the configura-

Assaf Harel, Department of Psychology, Hebrew University of Jerusalem, Jerusalem, Israel; Shlomo Bentin, Department of Psychology and Center of Neural Computation, Hebrew University of Jerusalem.

This study was funded by National Institute of Mental Health Grant R01 MH 64458 to Shlomo Bentin.

Correspondence concerning this article should be addressed to Assaf Harel, Department of Psychology, Hebrew University of Jerusalem, Jerusalem, 91905 Israel. E-mail: assaf.harel@mail.huji.ac.il

¹ Note, however, that the subordinate level in face categorization and subordinate level in object categorization were not equivalent in Tanaka's (2001) study. Whereas at the subordinate level for faces, stimuli were categorized on the basis of their individual identity (e.g., Bill Clinton), at the subordinate level for objects, stimuli were categorized on the basis of their species (e.g., robin), that is, on the basis of their membership in a group of exemplars.

² We suggest the term *canonical structure* as defined here instead of *global structure* as used by some authors in the literature (e.g., Bentin, Golland, Flevaris, Robertson, & Moscovitch, 2006) to avoid confusions with the term *global processing*, which entails holistic perception.

tion of a car's parts. Likewise, faces are defined by a canonical structure that is unique to them. This structure includes two eyes symmetrically located relative to an axis comprising the nose and the mouth (first-order relations; Maurer, Le Grand, & Mondloch, 2002), all placed within a roughly oval shape. The shared structure of exemplars within basic-level categories implies that within-category distinctions must be based on perceptual information that is more elaborate than that required for basic-level categorization, such as the precise description of the composing parts and/or their metric configuration (second-order relations; Maurer et al., 2002). Critically, the visual information required for making within-category, subordinate categorizations varies with the visual homogeneity of the subordinate groups, which defines their distinction. Moreover, the elaborate information, used for subordinate categorization, might also differ among different basic-level categories. It is this difference that might be domain specific to faces or, more generally, distinguish between perceptual processes applied to different object categories.

The differential use of visual information while processing objects and faces might be reflected in the spatial frequency (SF) scales used for basic and subordinate levels of categorization. Different SF scales convey different types of information: A high-pass filtered (HSF) image, for instance, preserves the sharp, fine-scale details of the original image. In contrast, a low-pass filtered (LSF) image retains the large-scale luminance variations, such as the light and dark blots and blobs in the image, which provide "a useful skeleton of the image" (Morrison & Schyns, 2001, p. 455) while fine details might be lost. Consequently, these two ranges of SFs have been traditionally associated with the division between local parts and their configuration, respectively (for a review, see Morrison & Schyns, 2001). In fact, spatial filtering of images has been used to explore the relative importance of parts and configuration during visual processing of faces and objects. Although this association between HSF and local processing and between LSF and configural processing has been challenged (Boutet, Collin, & Faubert, 2003), recent studies supported the above associations by manipulating SF and configural information conjunctively (Goffaux, Hault, Michel, Vuong, & Rossion, 2005; Goffaux & Rossion, 2006). Notwithstanding this debate, it is clear, however, that details are more difficult to discern in LSF than in HSF images. Therefore, at the very least, spatial filtering might be efficient to assess the necessity of details in basic-level and subordinate-level categorization of faces and objects.

In line with the above conceptualization, when categorization is based on the canonical structure of a category, the absence of details characteristic to LSF images should interfere less with performance as shape information is retained. Moreover, because at least the first-order relation of parts is retained in HSF images, the conceptualization of basic-level categorization described above should predict that it is not differentially affected by low- or high-spatial filtering (for a similar argument concerning object recognition, see Biederman & Kalocsai, 1997). By contrast, when elaborate information is required, low spatial frequencies might not be enough for efficient categorization. Accordingly, the level of performance with HSF images in subordinate categorization should indicate the necessity of details for subordinate categorization. Specifically, if details are required for subordinate categorization, then performance with LSF images should not be as good as with HSF images. Equal (or better) subordinate categorization perfor-

mance with LSF images should indicate that the type of elaborate information required by this process does not depend on details.

Support for a differential role of SFs at basic and subordinate levels of categorization was found by Collin and McMullen (2005) in a category verification paradigm with spatially filtered objects. In this experiment, there was almost no effect of SF on categorization performance either at the basic level (*dog*) or at the superordinate level (*animal*). In contrast, at a subordinate level (*collie*), categorization of LSF images was slower and less accurate than the categorization of either broadband (BB) or HSF images. These authors suggested that the vulnerability of subordinate category verification to low-pass filtering is due to its reliance on diagnostic details such as texture patterns or the parts' metric differences that are filtered out by low-pass filtering. Conversely, basic-level category verification is immune to spatial filtering, probably because it relies on shape information that is retained in both low- and high-frequency scales.

It is possible, however, that Collin and McMullen's (2005) conclusions are limited to the object categories that they used (vehicles and animals). In particular, the generalization of these findings to face processing might be unwarranted because, as suggested above, subordinate categorization of faces might rely on diagnostic information that is different than that used for subordinate categorization of objects. Specifically, it has been shown that subordinate distinctions between faces (e.g., race or gender) require configural information (e.g., Rhodes, Tan, Brake, & Taylor, 1989, and Campanella, Chrysochoos, & Bruyer, 2001, respectively). This type of information, however, might not be required for subordinate categorization of cars or airplanes (Biederman, Subramaniam, Bar, Kalocsai, & Fiser, 1999). Therefore, it is possible that filtering out low SFs from the image would harm subordinate face processing more than it would harm objects. Many studies have looked into the role of specific SF bands in face recognition (for a review, see Ruiz-Soler & Beltran, 2006), but only a few have directly contrasted the use of spatial frequencies in recognition of faces and objects (Biederman & Kalocsai, 1997; Boutet, Collin, & Faubert, 2003; Collin, Liu, Troje, McMullen, & Chaudhuri, 2004; Goffaux, Gauthier, & Rossion, 2003; Parker, Lishman, & Hughes, 1996). Although the findings were not fully consistent, all of these studies reported a difference between the frequency scales used for processing faces and objects. However, it is not clear where these differences stem from. Specifically, because all of these studies did not control for the level of representation accessed during processing, it might be possible that the differences between faces and objects stem from different default levels of categorization, each relying on different perceptual information.

The category verification task, introduced by Rosch and colleagues (Rosch et al., 1976), solves the problem of uncontrolled level of representation to some extent. In this paradigm, a category label is presented first, followed by a picture of an object. The task is to indicate whether the object is a member of the category. Although this task obviously includes perceptual semantics elicited by the verbal label, using labels of varying degrees of specificity (superordinate, basic, subordinate) allows the control of categorization level while controlling the visual appearance of the object by either keeping it constant or manipulating it to assess the influence of perceptual factors on the categorization of the object.

Manipulating the level of categorization for objects as well as for faces while they are spatially filtered might elucidate the interrelations between levels of categorization, object category, and the perceptual information (as contained in different SF scales) supporting object

categorization. All three variables are hypothesized to influence categorization performance, but they are not of equal interest in the present context. Specifically, SF scale preferences were considered here not as an end in themselves but as a measure to estimate the diagnostic information that is needed for object categorization. Thus, the interaction between SF and the task-imposed categorization level on the one hand and between SF and stimulus category on the other hand may provide an answer to how top-down categorization and bottom-up processing of the image influence the utilization of information supporting object recognition. For example, an outcome showing that the task-imposed level of categorization determines the use of SF information across categories would support the idea of a flexible usage of diagnostic information according to which the SF information used during perceptual processing is modulated by top-down factors, such as the categorization task (Schyns, 1998; Schyns & Oliva, 1999; Sowden & Schyns, 2006). Similarly, finding different effects of SF on different categories irrespective of the top-down imposed categorization level might reflect a strongly imposed default level of categorization that might be different for different stimulus categories. Finally, finding different patterns of spatial filtering effects on subordinate and basic levels of categorization for faces and objects (i.e., a second-order interaction between SF, level of categorization, and stimulus category) would indicate flexible usage of SF scales that is nevertheless influenced by the stimulus category.

In the present study, we assessed performance in a category verification task to explore the complex interactions between top-down-determined recognition goals (i.e., levels of categorization) and the perceptual information (SF) available in images of faces, cars, and airplanes. Assuming that basic-level categorization does not require detail information, we hypothesized that there would be no difference in the utilization of SF between faces and objects while determining their basic-level category. Specifically, on the basis of Collin and McMullen's (2005) findings, we predicted that low-pass filtering and high-pass filtering would affect basic-level categorization to the same degree. In contrast, because the diagnostic information needed for subordinate categorization of faces and objects is probably different (and might also depend on the particular subordinate distinction required; Johnson & Mervis, 1997), we predicted that spatial filtering would influence differently the subordinate categorization of faces and objects. In particular, assuming that subordinate distinctions within airplanes and within cars are based on details, they should rely to a great extent on high spatial frequencies. In contrast, because subordinate categorizations of faces are based, at least partly, on configural computations, we predicted that filtering out low spatial frequencies would reduce subordinate categorization performance for faces.

Method

Participants

Twenty-two students from the Hebrew University of Jerusalem (11 men; mean age 25 years) participated in the experiment as part of course requirements or for monetary compensation (60 new Israeli shekels, approximately \$15). All had normal or corrected-to-normal visual acuity and no history of psychiatric or neurological disorders. Participants signed an informed written consent form according to the requirements of the institutional review board of the Hebrew University of Jerusalem.

Stimuli

The original stimuli consisted of 80 different images of female faces in front view, half being Chinese and half being Israeli (Caucasian); 80 different images of cars in side view, half of European makers and half of Japanese makers; and 80 different images of airplanes in side view, half being combat jets and half being civil airliners. Face images were selected from our database, whereas images of cars and airplanes were downloaded from the Internet. All images were 360×360 pixels, which, seen from a distance of 70 cm, subtended 9.9° of visual angle. The object size was equated across categories by scaling each stimulus within the image to an identical diagonal of 370 pixels within its bounding box (excluding the background; cf. Collin & McMullen, 2005). The resizing preserved the aspect ratio of the stimuli, which varied across the different object categories. Mean luminance and root-mean-square contrast were equated across images and categories (not including the gray background in calculation) using Adobe Photoshop. The background of all of the objects was a uniform gray equated to the mean object luminance.

The original broadband (BB) images were spatially filtered in MATLAB (Version 7.0.1; <http://www.mathworks.com>) using a Butterworth filter with an exponent of 4. The LSF and HSF cutoff corners were 1 cycle/degree and 6.5 cycles/degree, respectively. These values were set to match those used in previous studies (e.g., Goffaux et al., 2003), and, given the size of our stimuli, they correspond to approximately 10 cycles/image and 65 cycles/image for the LSF and HSF conditions, respectively. Altogether, there were 720 different images, 240 in each spatial filter condition (BB, HSF, and LSF). Examples of all stimulus conditions may be seen online in supplemental materials.

Design and Procedure

Participants performed a category verification task in two consecutive sessions. A trial started with the presentation of a Hebrew³ object category label presented for 500 ms at the center of the screen using 24-point Courier New font. The label was followed by a fixation cross presented for 250 ms, an image of an object presented for 300 ms, and then a screen that remained blank until the participant responded. Randomly selected intertrial intervals of 500, 800, 1,000, or 2,000 ms separated the next trial from the response.⁴ Participants had to indicate by pressing one of two buttons whether the object matched the category label or not, and they were instructed to respond as quickly as possible but avoid making mistakes. Accuracy rates and reaction times (RTs) were recorded. RTs were acquired from the onset of the stimulus.

Basic-level (face, car, and airplane) and subordinate-level (Chinese or Israeli face, European or Japanese car, civil airliner or combat jet) category labels were presented in two separate sessions using the same stimuli. The order of the sessions was counterbalanced across participants. A session consisted of five blocks, each consisting of 144 trials. The nine experimental conditions (three categories and three SF scales) were mixed in each block and presented in random

³ There are no different uppercase and lowercase sets in Hebrew.

⁴ The different intertrial intervals were needed for an event-related potentials study that was done in parallel.

order. In half of the trials in each block, the image corresponded with the preceding category label (match trials), and in the other half, the image did not correspond with the preceding label (mismatch trials). At the subordinate categorization level, images in the mismatch trials were from the same basic-level category as the category label. For example, the category label *Chinese face* was followed by an image of an Israeli face and the category label *combat jet* was followed by an image of an airliner. In the basic-level condition, mismatch images were from the two object categories other than the object category label. For example, following the label *car*, an image of an airplane or an image of a face appeared with equal probability. No stimulus repeated itself within a block. Each session of the experiment was preceded by instructions and a training session of 72 trials that presented all of the experimental conditions in equal proportion. After they completed the training session, participants were provided with feedback on their performance and training was repeated if necessary.

Results

Subordinate categorization of cars was at chance regardless of the SF scale (53.1%, 50.8%, and 50.7% correct for the BB, HSF, and LSF conditions, respectively). Consequently, we excluded the car category from all further statistical analyses. Furthermore, as in all previous studies (e.g., Collin & McMullen, 2005; Hamm & McMullen, 1998; Johnson & Mervis, 1997; Scott, Tanaka, Sheinberg, & Curran, 2006; Tanaka, 2001; Tanaka & Taylor, 1991), basic-level accuracy was very high. It is interesting that this high level of accuracy was observed in the present study across all SF scales. Therefore, our conclusions are based primarily on RT patterns, and we report the accuracy analysis only as supporting data.

A repeated-measures analysis of variance (ANOVA) was used to analyze the RTs and the accuracy in each experimental condition. The factors were level of categorization (basic, subordinate), category (faces, airplanes) and spatial frequency scale (BB, LSF, HSF).⁵ For factors with more than two levels, the degrees of freedom were corrected for nonsphericity using the Greenhouse-Geisser adjustment (for simplicity, the uncorrected degrees of freedom are presented).

RT Results

Mean RTs for the different experimental conditions are presented in Figure 1. There was no significant main effect of category, $F(1, 21) = 1.14$, $MSE = 27,351$, $p = .30$, partial $\eta^2 = .05$. Basic-level categorization was significantly faster than subordinate-level categorization across category and SF scale, $M_{\text{basic}} = 517$ ms, standard error of the mean (SEM) = 26 ms, $M_{\text{subordinate}} = 710$ ms, $SEM = 35$ ms; $F(1, 21) = 44.27$, $MSE = 55,711$, $p < .001$, partial $\eta^2 = .68$. Particularly worth noticing is the absence of a Category \times Level of Categorization interaction, $F(1, 21) = 1.36$, $MSE = 29,573$, $p = .25$, partial $\eta^2 = .06$. The main effect of spatial frequency scale on speed of categorization was significant, $F(2, 42) = 14.44$, $MSE = 1,484$, $p < .001$, partial $\eta^2 = .41$, but qualified by a significant interaction with category, $F(2, 42) = 17.63$, $MSE = 874$, $p < .001$, partial $\eta^2 = .46$, as well as with level of categorization, $F(2, 42) = 4.71$, $MSE = 908$, $p < .02$, partial $\eta^2 = .18$.

Critically, there was a significant three-way interaction of category, level of categorization, and SF scale, $F(2, 42) = 8.27$, $MSE = 865$, $p < .001$, partial $\eta^2 = .28$. The analysis of simple

effects showed that the utilization of SF was different for airplanes and faces. Basic-level categorization of airplanes, $F(2, 42) = 10.27$, $MSE = 440$, $p < .001$, partial $\eta^2 = .33$, was significantly faster in the BB condition than in both the HSF (Bonferroni corrected, $p < .04$) and LSF (Bonferroni corrected, $p < .04$) conditions, which did not differ (Bonferroni corrected, $p > .05$). Subordinate categorization of airplanes, $F(2, 42) = 25.01$, $MSE = 1,702$, $p < .001$, partial $\eta^2 = .54$, however, was delayed in the LSF condition relative to categorization in the BB condition (~ 80 ms; Bonferroni corrected, $p < .001$) four times more than in the HSF condition relative to the BB condition (~ 20 ms; Bonferroni corrected, $p < .05$). RTs were significantly longer in the LSF condition than the HSF condition (Bonferroni corrected, $p < .001$). Faces, however, showed a different pattern. Basic-level categorization of faces, $F(2, 42) = 16.32$, $MSE = 587$, $p < .001$, partial $\eta^2 = .43$, was delayed by HSF relative to BB (Bonferroni corrected, $p < .01$) and relative to LSF (Bonferroni corrected, $p < .001$) without further distinction between BB and LSF (Bonferroni corrected, $p > .05$). However, at the subordinate level, $F(2, 42) = 25.36$, $MSE = 3,260$, $p < .001$, partial $\eta^2 = .54$, both HSF and LSF conditions delayed categorization of faces relative to the BB condition (Bonferroni corrected, $p < .001$ and $p < .007$, respectively), although this delay was more than five times larger for HSF images (~ 87 ms) than for LSF images (~ 17 ms).

Finally, to explore possible differential effects of the SF scale on each stimulus group at the subordinate level, we compared the RTs to Chinese and Israeli faces and to combat jets and civil airlines in separate SF Scale \times Subcategory ANOVAs (Table 1 and Table 2). For faces, ANOVAs showed a main effect of SF scale, $F(2, 42) = 36.52$, $MSE = 5,076$, $p < .001$, partial $\eta^2 = .63$; no main effect of subcategory, $F(1, 21) < 1.00$, partial $\eta^2 = .03$; and an interaction that approached significance, $F(2, 42) = 3.45$, $MSE = 3,484$, $p = .056$, partial $\eta^2 = .14$. As evident in Table 1, although for both Chinese and Israeli faces, the RTs were faster in the LSF than the HSF conditions, the difference was twice as big for Chinese than for Israeli faces, $F(1, 21) = 5.10$, $MSE = 6,777$, $p < .04$, partial $\eta^2 = .19$.

For airplanes, ANOVAs showed a main effect of SF scale, $F(2, 42) = 21.65$, $MSE = 4,332$, $p < .001$, partial $\eta^2 = .51$; a marginal main effect of subcategory, $F(1, 21) = 4.00$, $MSE = 6,249$, $p = .059$, partial $\eta^2 = .16$; and a significant interaction, $F(2, 42) = 4.07$, $MSE = 3,638$, $p < .05$, partial $\eta^2 = .16$. As evident in Table 2, although for both combat jets and civil airliners, the RTs were faster in the HSF than the LSF conditions, the difference was about four times bigger for airliners than for combat jets, $F(1, 21) = 22.78$, $MSE = 7,266$, $p < .001$, partial $\eta^2 = .52$.

⁵ Analyzing the RTs and accuracy with the addition of the response factor (match, mismatch) showed a significant or main effect of response that approached significance, $F(1, 21) = 85.40$, $MSE = 5,911$, $p < .001$, and $F(1, 21) = 4.00$, $MSE = 0.01$, $p = .06$, for RTs and accuracy, respectively. These main effects reflected faster and more accurate RTs in the match than in the mismatch response conditions. Because the response factor did not have any differential effect on the second-order interaction between SF, level of categorization, and stimulus category, which is the focus of this study, either for RTs, $F(2, 42) = 1.42$, $MSE = 1,075$, $p = .25$, or for accuracy, $F(2, 42) = 2.33$, $MSE = 0.003$, $p = .11$, all further analyses were carried out with match and mismatch response conditions collapsed.

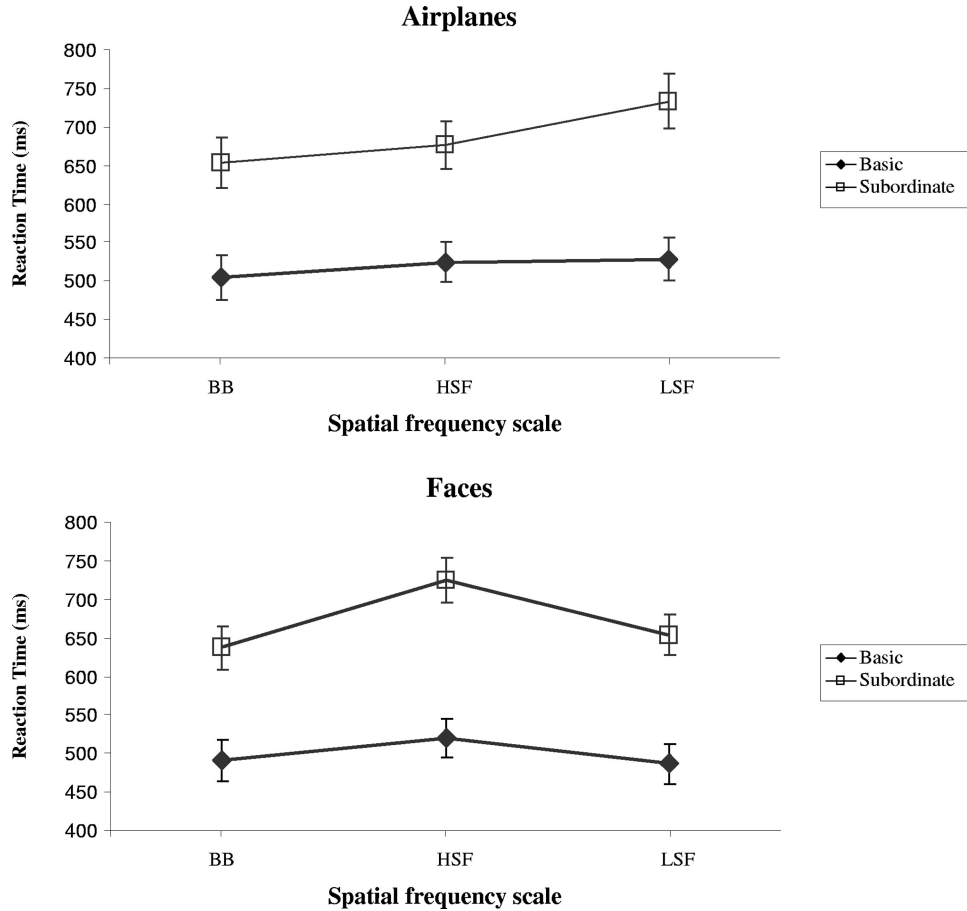


Figure 1. Mean reaction times for categorization of airplanes and faces at the basic level and subordinate level presented in the different spatial frequency scales (broadband [BB], high-pass filtered [HSF], low-pass filtered [LSF]). Error bars represent standard errors of the means.

Accuracy Results

The percentage of correct responses in the different experimental conditions is presented in Figure 2. As is evident in that figure, the pattern of accuracy mirrored the RT results. Basic-level categorization was more accurate (94%) than subordinate categorization (83%) across object categories and SF scales, $F(1, 21) = 53.08, MSE = 142, p < .001, \text{partial } \eta^2 = .72$. Accuracy rates did

not differ as a function of category, $F(1, 21) < 1.00$, and there was no significant interaction between category and categorization level, $F(1, 21) = 1.83, MSE = 109, p = .19, \text{partial } \eta^2 = .08$. SF scale, however, did affect accuracy rates, $F(2, 42) = 4.58, MSE = 352, p < .02, \text{partial } \eta^2 = .18$, and this effect was qualified by significant interactions with category, $F(2, 42) = 23.91, MSE = 1,628, p < .001, \text{partial } \eta^2 = .53$, and level of categorization, $F(2, 42) = 3.84, MSE = 2,957, p < .04, \text{partial } \eta^2 = .15$. Further

Table 1
Reaction Times (RTs) in Milliseconds and Standard Errors of the Mean (SEMs) Within the Face Subordinate-Level Condition

Subcategory (race)	Spatial frequency scale					
	BB		HSF		LSF	
	RT	SEM	RT	SEM	RT	SEM
Israeli	636	26	719	30	668	29
Chinese	638	28	762	36	655	27

Note. BB = broadband; HSF = high-pass filtered; LSF = Low-pass filtered.

Table 2
Reaction Times (RTs) in Milliseconds and Standard Errors of the Mean (SEMs) Within the Airplane Subordinate-Level Condition

Subcategory (type)	Spatial frequency scale					
	BB		HSF		LSF	
	RT	SEM	RT	SEM	RT	SEM
Civil airliner	668	35	674	32	763	41
Combat jet	642	33	674	28	707	34

Note. BB = broadband; HSF = high-pass filtered; LSF = low-pass filtered.

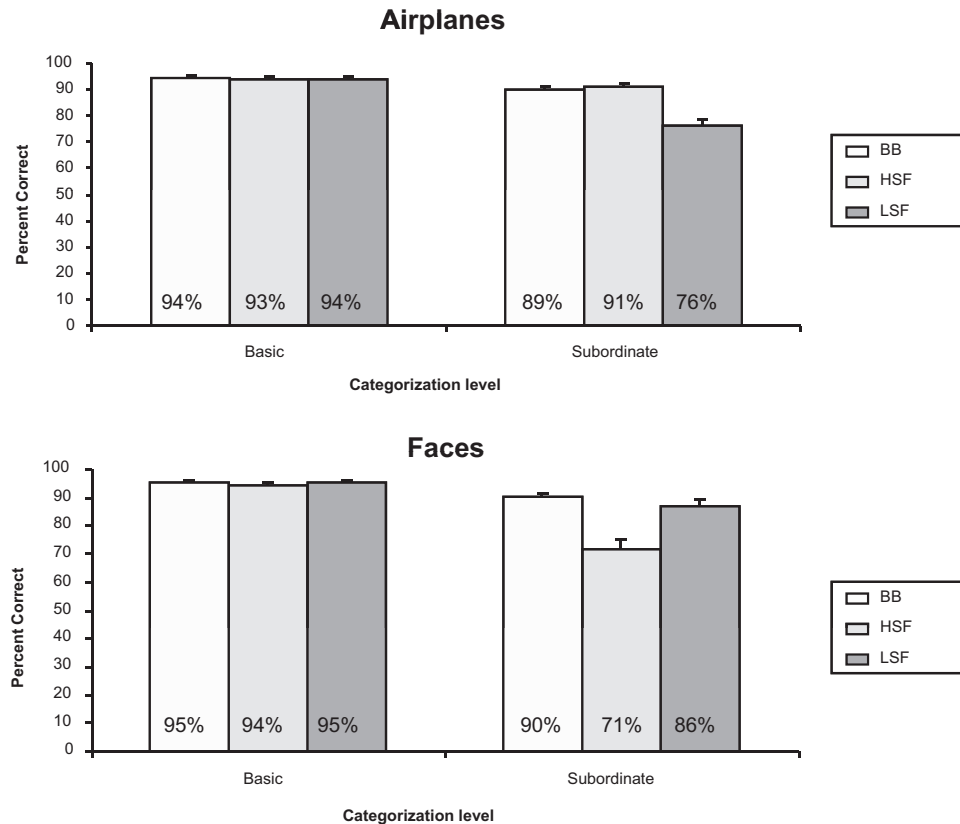


Figure 2. Mean percentage of accuracy for airplanes and faces at the basic level and subordinate level of categorization presented in the different spatial frequency scales (broadband [BB], high-pass filtered [HSF], low-pass filtered [LSF]).

analysis of the SF Scale \times Category interaction revealed a significant SF effect for both airplanes, $F(2, 42) = 5.85$, $MSE = 413$, $p < .02$, partial $\eta^2 = .22$, and faces, $F(2, 42) = 26.82$, $MSE = 844$, $p < .001$, partial $\eta^2 = .56$. Post hoc analysis of simple effects showed that for airplanes, across categorization level, accuracy rates in the BB condition were similar to rates in the HSF condition (Bonferroni corrected, $p > .05$), which in turn were significantly higher than rates in the LSF condition (Bonferroni corrected, $p < .01$). In contrast, in faces, accuracy rates were equivalent in the BB and LSF conditions (Bonferroni corrected, $p > .05$), with both of them significantly higher than rates in the HSF condition (for the BB condition, Bonferroni corrected, $p < .001$; for the LSF condition, $p < .001$). An analysis of the interaction between SF scale and categorization level revealed a significant SF effect at the subordinate-but not at the basic-level categorization: For subordinate-level categorization, $F(2, 42) = 4.42$, $MSE = 343$, $p < .02$, partial $\eta^2 = .17$; for basic-level categorization, $F(2, 42) = 1.00$. Post hoc analyses revealed that accuracy rates in the subordinate level were higher in the BB condition compared with the HSF condition (Bonferroni corrected, $p < .03$), which did not differ from the LSF condition (Bonferroni corrected, $p > .05$).

Like for RTs, of special importance for the goals of the present study was finding a significant three-way interaction, $F(2, 42) = 20.88$, $MSE = 74$, $p < .0001$, partial $\eta^2 = .50$. Subsequent analysis of this interaction revealed that at the basic level, the SF

scale had no effect on accuracy regardless of category: For airplanes, $F(2, 42) = 1.20$, $MSE = 4$, $p = .30$, partial $\eta^2 = .05$; for faces, $F(2, 42) < 1.00$. In contrast, SF scale affected subordinate categorization of each category differently. For airplanes, the significant effect of SF scale, $F(2, 42) = 6.33$, $MSE = 287$, $p = .01$, partial $\eta^2 = .23$, reflected lower accuracy rates for LSF images than for HSF images (Bonferroni corrected, $p < .001$). An opposite effect was found for faces: The significant effect of SF scale, $F(2, 42) = 27.59$, $MSE = 108$, $p < .001$, partial $\eta^2 = .57$, reflected lower accuracy for the HSF images than for LSF images (Bonferroni corrected, $p < .001$) or BB images (Bonferroni corrected, $p < .001$), which did not differ (Bonferroni corrected, $p > .05$).

Finally, a signal detection approach was used to assess possible response biases during the subordinate categorization process (cf. Gauthier & Bukach, 2007). Although the major purpose of this analysis was to assess possible changes in the criterion across different SF scales, we calculated and analyzed both d' and the criteria (Table 3). The d' and the criteria were calculated by collating the responses of the participants for each trial (i.e., match or mismatch between the category label and the image) and comparing the results with the required correct responses (match or mismatch between the category label and the image). Given the nature of the category verification task, the two measures were calculated and analyzed for each subordinate category in a given SF scale separately. A multiple t test (Bonferroni-corrected) pro-

Table 3
*Mean d' and Criterion Values Within the Face
 Subordinate-Level Condition and Within the Airplane
 Subordinate-Level Condition*

Subcategory	Spatial frequency scale					
	BB		HSF		LSF	
	d'	Criterion	d'	Criterion	d'	Criterion
Race						
Israeli	2.74	.18	2.20	.07	2.45	.07
Chinese	2.69	-.01	0.81	.17	2.43	.01
Type						
Civil airliner	2.70	.17	2.73	.14	1.27	.01
Combat jet	2.70	.17	2.83	.13	2.07	.01

Note. BB = broadband; HSF = high-pass filtered; LSF = low-pass filtered.

cedure was applied to verify whether any of the criteria differed significantly from zero (to test for bias). This analysis revealed that none of the 12 criteria differed significantly from zero, indicating that no strategic biases have been applied during verification of all subordinate categories. The analysis of d' complemented the pattern of the results presented above, adding discrimination accuracy values within each subordinate category. Separate one-way ANOVAs followed by univariate contrasts for each subordinate category with SF as independent variable showed that d' was not affected by SF for Israeli faces, $F(2, 42) = 2.13$, $MSE = 0.8$, $p = .13$, partial $\eta^2 = .09$, whereas for Chinese faces, d' was significantly lower for HSF images than for both BB and LSF images, $F(2, 42) = 22.80$, $MSE = 1.55$, $p < .001$, partial $\eta^2 = .52$. For airplanes, this analysis showed that performance with both combat jets and airliners was reduced in LSF relative to both the BB and HSF conditions, $F(2, 42) = 12.40$, $MSE = 0.32$, $p < .001$, partial $\eta^2 = .37$, and $F(2, 42) = 36.66$, $MSE = 0.48$, $p < .001$, partial $\eta^2 = .63$, respectively.

Discussion

In the present study, we used spatial filtering to index the type of information extracted from images of faces and objects for basic-level and subordinate-level categorization. Our goal was to examine how different levels of categorization affect the processing of faces and nonface object categories and whether these categories differ in the type of visual information needed for categorization at basic and subordinate levels. From these data, we sought to determine whether the type of visual representation that contacts the semantic system initially (the entry point; Jolicoeur et al., 1984) is task dependent and flexible, whether this flexibility is different for faces and objects, or whether the entry point is rigidly imposed by the category type.

Across SF conditions, subordinate categorization of airplanes as well as faces was delayed and less accurate relative to basic-level categorization. In addition, for both stimulus categories, performance was superior in the BB condition than when spatial filters were applied. It is important to note, however, that the pattern of spatial filtering effects on categorizing faces and nonface objects differed as a function of categorization level. At the basic level,

categorization of airplanes was equally delayed for high- and low-pass filters. In contrast, for faces, low-pass filtering did not significantly delay performance relative to performance in the BB condition. At the subordinate level, the absence of high spatial frequencies (LSF condition) impaired categorization of airplanes to a greater degree than did the absence of low spatial frequencies (HSF condition). For faces, the inverse effect was found: Performance was not as good in the HSF condition compared with the LSF condition.

Recall that the nature of diagnostic information used in each condition was indexed in the present study by the different SF scales. Thus, the second-order interaction between SF scale, level of categorization, and stimulus category suggests that the diagnostic information used for visual categorization is neither purely category determined nor solely determined top-down by the categorization task. This interpretation challenges models assuming that the spatial frequencies used during image processing are predominantly determined either by the stimulus characteristics or by the task demands. For example, ignoring possible task effects, some authors suggested that the SF scales used during face processing differ from those used during letter processing (Gold, Bennett, & Sekuler, 1999). Similarly, other authors suggested that faces are sensitive to SF manipulations, whereas objects are not (Biederman & Kalocsai, 1997; Collin et al., 2004). The alternative view, stressing the predominance of the task, has also been proposed. For example, Schyns and his colleagues suggested that the SF scales extracted from an image are determined by conceptual top-down factors, such as the particular categorization task (Schyns, 1998; Schyns, Bonnar, & Gosselin, 2002; Schyns & Oliva, 1999; for reviews, see Morrison & Schyns, 2001; Sowden & Schyns, 2006).

In contrast to these two alternatives, the present data suggest that visual categorization is constrained by both stimulus and task factors. Specifically, we found that different SF scales are essential for processing different categories, but this effect is modulated by the level of categorization required. Whereas for faces, low spatial frequencies were needed more than high spatial frequencies at basic as well as subordinate levels, for airplanes, both frequency scales were equally needed at the basic level, whereas high frequencies were more important than the low frequencies at the subordinate level.

Although, as mentioned in the introduction, the existence of a simple relationship between low- and high-frequency scales and global and local processing of visual information is controversial (Boutet et al., 2003; Goffaux et al., 2005; Goffaux & Rossion, 2006; Loftus & Harley, 2004; Morrison & Schyns, 2001; Sowden & Schyns, 2006), details are evidently absent from LSF images, whereas the configural information as well as information about texture, although they might be present, are more difficult to discern when the low spatial frequencies are removed (i.e., the HSF condition). Therefore, the present data suggest that whereas subordinate categorization of airplanes requires detailed information about diagnostic parts, for faces, the absence of details (the LSF condition) had only a small effect and, it is important to note, this effect was similar for basic and subordinate levels of categorization.

The importance of canonical structure for basic-level categorization of faces is in line with previous studies proposing that face detection is based on their global shape (e.g., Bentin et al., 2006). However, the apparent importance of canonical structure as well as

configural information for race categorization is more intriguing. Previous studies suggested that holistic⁶ (Michel, Caldara, & Rossion, 2006; Tanaka, Kiefer, & Bukach, 2004) as well as configural processing (Rhodes, Hayward, & Winkler, 2006; Rhodes et al., 1989) is more prevalent for own-race than other-race faces. Because our participants were Israelis, we expected that the absence of low frequencies (which are assumed to better preserve the second-order relations between parts) would slow the RTs more for Israeli faces than for Chinese faces. Surprisingly, we found an inverse tendency. As supported by the Subcategory \times SF interaction, relative to the RTs in the BB condition, the RTs in the HSF condition were delayed twice as much as those in the LSF condition when the faces were Chinese than when they were Israeli. This pattern suggests that, although configural information might be important for categorization of both races, impeding configural processing impairs race categorization of other-race faces to a greater extent than it impairs categorization of own-race faces. However, note that all of the studies that found a configural advantage for own-race faces focused on the *other-race effect*, that is, the faster and more accurate recognition of individual own-race than other-race faces. In contrast, in the present study, faces were categorized by race rather than as individual exemplars, a categorization that is actually faster for other-race than own-race faces (the *other-race advantage*; Caldara, Rossion, Bovet, & Hauert, 2004; Levin, 1996; Valentine & Endo, 1992; Zhao & Bentin, in press). We are not aware of any previous study of the importance of configural information for subordinate categorization of faces by race. Although not conclusive, a possible interpretation of the present data is that although for both own- and other-race faces the recognition of a face's race relies primarily on configural processes and texture, the own-race face could be more easily recognized than the other-race face on the basis of details provided in the high-frequency scale.

It is interesting to note that the interaction between the type of airplane and SF scale was also significant. In the absence of previous studies exploring either the role of spatial frequencies or the relative importance of details for subordinate categorization of objects, our account for this effect is by necessity post hoc. The interaction between SF scales and type of airplane revealed that although for both types of airplanes, the absence of high frequencies delayed the RTs relative to the RTs in the BB condition, this delay was four times bigger for civil airliners than for combat jets. Moreover, for civil airliners, the RTs in the HSF condition were similar to those in the BB condition. This suggests that the absence of low frequencies had no effect on the categorization of civil airliners. In contrast, the subordinate categorization of combat jets was 32 ms faster in the BB condition than in the HSF condition. Together, these results suggest that the categorization of civil airliners was based heavily on details, whereas the categorization of combat jets was based on both details and configuration. A glimpse at the (representative) example shown in the supplemental materials demonstrates that a conspicuous feature of airliners is the line of windows along their body, which are retained in the HSF condition but are indiscernible in the LSF condition. No such conspicuous detail characterizes combat jets, which may explain the need for the additional information contained in low spatial frequencies. The different reliance of combat jets and airliners on details implies that, unlike faces, which have a more homogeneous structure, the subcategorization of objects is more flexibly deter-

mined by the distinctive features of each subordinate group. In line with this assumption, by parametrically manipulating image HSF information, previous studies have shown that different nonface object categories recruit information about parts to a different degree (Vannucci, Viggiano, & Argenti, 2001; Viggiano, Constantini, Vannucci, & Righi, 2004). In contrast to subordinate categorization, basic-level categorization can be (and probably is) performed on the basis of the canonical structure defining the category. This structure is retained by both high and low SF scales (Collin & McMullen, 2005).

A caveat to the above interpretation is the possibility that the SF effect on subordinate categorization of faces and objects reflects differences in response bias rather than differences in the ability to discriminate between the subordinate categories (Gauthier & Bukach, 2007; Richler, Gauthier, Wenger, & Palmeri, 2008). This caveat, however, is dismissed by the signal detection analysis of the subordinate performance. Whereas the *d*'s were in line with the accuracy pattern, the response criterions in all conditions were not different from zero, indicating that no strategic bias affected performance.

An unexpected finding in this study was that across filters, the level of categorization had similar effects on faces and airplanes. Specifically, subordinate categorization for faces as well as for airplanes was slower and less accurate. Whereas slower subordinate-than basic-level categorization may be expected for objects (Rosch et al., 1976), previous data suggested that faces are categorized equally fast at basic and subordinate levels (Tanaka, 2001). Note, however, that in Tanaka's (2001) study, the subordinate categorization was at the individual exemplar level using highly familiar, distinctive faces. A more recent study from the same group showed that category verification of individual faces takes longer than basic-level categorization when the distinctiveness of the faces is reduced (D'Lauro, Tanaka, & Curran, 2008). In addition to using unfamiliar faces, in the present study, we did not have the subordinate level of categorization at the individual level. The different subordinate levels of categorization addressed in the two studies and the different familiarity with the faces could account for the longer RTs in the subordinate- relative to basic-level categorization found here compared with the absence of such effect in Tanaka's study. If the goal of the face processing system is to determine the face identity (either by default or task determined), the diagnostic information that is initially extracted from the face image should address face identity rather than its race. Consequently, race categorization would wait for the relevant diagnostic information to be extracted later. This strategy could be particularly true for familiar faces, which have a preexistent unique representation.

Still additional considerations are required to explain why race categorization takes longer than basic-level categorization. Assuming that the entry point for faces is at the individual level and race categorization addresses an intermediate level between individual and basic, a straightforward hierarchical model should predict faster RTs for race categorization than for basic categorization. Because this was not the case, the present data suggest that the process of face categorization—in contrast to object categorization—is not hierarchical, in the sense that observers do not identify

⁶ The term *holistic* was used in these studies to describe the integrative aspect of processing face parts into a gestalt.

a face in a serial process of reducing the category size sequentially. Rather, concurrent with the detection of a face (basic-level categorization), the identity-relevant information is extracted while information relevant to larger subordinate grouping is accumulated at different rates. That is, George Bush is identified separately (and perhaps prior) to identifying his Caucasian race, his masculinity, or his middle age.

It is possible, in fact, that categorization hierarchies are irrelevant to faces. Note, indeed, that Bruce and Young's (1986) classic model of face perception has no hierarchical links between the recognition of gender or age (subordinate categorization that is achieved via directed visual processes) and the activation of face recognition units. Although both processes entail an initial view-centered description during the structural encoding, after this initial stage, individual face identification and less specific subordinate categorization (e.g., race in the present study) follow separate paths (and different processing rates). This nonhierarchical category structure might, however, be peculiar to faces, whereas for objects, the hierarchical organization is eminent. If this is correct, perhaps there is no good reason to compare subordinate categorization of faces with that of other objects (see also Biederman et al., 1999), including objects of expertise (for a discussion, see Bukach, Gauthier, & Tarr, 2006). Regarding expertise, although it stands to reason that, like for faces, the entry point of objects of expertise is at the individual level (Tanaka & Gauthier, 1997; Tanaka & Taylor, 1991; Tarr & Gauthier, 2000), the process by which the subordinate categorization is achieved might be different. Because in the present study we did not examine expertise for nonface categories, the resolution of this question awaits further investigation.

In conclusion, using a within-subject comparison of category verification of faces and nonface objects at basic and subordinate levels of categorization while manipulating the spatial frequencies available in the image, we support a view claiming that both the task-determined level of categorization and the type of stimulus category influence the nature of visual information extracted during recognition. Assuming that details are missing in low SF scales whereas configural processes are impaired in high SF scales, the present data indicate that the configural structure contained in low frequencies is necessary for successful categorization of faces at both basic and subordinate levels, whereas details are not essential. In contrast, for objects, although basic-level categorization is based primarily on their canonical structure, the visual information needed for subordinate-level categorization varies across object categories reflecting the diagnostic features specific to each subcategory. In addition, the present data suggest that hierarchical structure of categorization does not apply equally for faces and objects. Specifically, the current study suggests that the entry point for visual recognition is flexible and determined conjointly by the stimulus category and the level of categorization, which reflects the observer's recognition goal.

References

- Bentin, S., Golland, Y., Flevaris, A., Robertson, L. C., & Moscovitch, M. (2006). Processing the trees and the forest during initial stages of face perception: Electrophysiological evidence. *Journal of Cognitive Neuroscience*, *18*, 1406–1421.
- Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society London: Biological Sciences*, *352*, 1203–1219.
- Biederman, I., Subramaniam, S., Bar, M., Kalocsai, P., & Fiser, J. (1999). Subordinate-level object classification reexamined. *Psychological Research*, *62*, 131–153.
- Boutet, I., Collin, C. A., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception and Psychophysics*, *65*, 1078–1093.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.
- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences*, *10*, 159–166.
- Caldara, R., Rossion, B., Bovet, P., & Hauert, C. A. (2004). Event-related potentials and time course of the 'other-race' face classification advantage. *NeuroReport*, *15*, 905–910.
- Campanella, S., Chrysochoos, A., & Bruyer, R. (2001). Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor. *Visual Cognition*, *8*, 237–262.
- Collin, C. A., Liu, C. H., Troje, N., McMullen, P., & 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.
- Collin, C. A., & McMullen, P. A. (2005). Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level categorization. *Perception and Psychophysics*, *67*, 354–364.
- D'Lauro, C., Tanaka, J. W., & Curran, T. (2008). The preferred level of face categorization depends on discriminability. *Psychonomic Bulletin and Review*, *15*, 623–629.
- Gauthier, I. (2000). What constrains the organization of the ventral temporal cortex? *Trends in Cognitive Sciences*, *4*, 1–2.
- Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition*, *103*, 322–330.
- Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Cognitive Brain Research*, *16*, 416–424.
- Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, *34*, 77–86.
- Goffaux, V., & Rossion, B. (2006). Faces are "spatial"—Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1023–1039.
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, *39*, 3537–3560.
- Hamm, J. P., & McMullen, P. A. (1998). Effects of orientation on the identification of rotated objects depend on the level of identity. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 413–426.
- Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General*, *126*, 248–277.
- Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, *16*, 243–275.
- Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1364–1382.
- Loftus, G. R., & Harley, E. M. (2004). How different spatial-frequency components contribute to visual information acquisition. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 104–118.
- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Science*, *6*, 255–260.
- Michel, C., Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition*, *14*, 55–73.
- Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the

- categorization of faces, objects, and scenes. *Psychonomic Bulletin and Review*, 8, 454–469.
- Parker, D. M., Lishman, J. R., & Hughes, J. (1996). Role of coarse and fine spatial information in face and object processing. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1448–1466.
- Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin and Review*, 13, 499–505.
- Rhodes, G., Tan, S., Brake, S., & Taylor, K. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, 80, 313–331.
- Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 328–342.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Bream, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.
- Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research*, 70, 273–292.
- Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. *Cognition*, 67, 147–179.
- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13, 402–409.
- Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, 69, 243–265.
- Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience*, 18, 1–13.
- Sowden, P. T., & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Sciences*, 10, 538–545.
- Tanaka, J. W. (2001). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General*, 130, 534–543.
- Tanaka, J. W., & Gauthier, I. (1997). Expertise in object and face recognition. *The Psychology of Learning and Motivation*, 36, 83–125.
- Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, 93, B1–B9.
- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23, 457–482.
- Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, 3, 764–769.
- Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113, 169–193.
- Valentine, T., & Endo, M. (1992). Toward an exemplar model of face processing: The effects of race and distinctiveness. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 44(A), 161–204.
- Vannucci, M., Viggiano, M. P., & Argenti, F. (2001). Identification of spatially filtered stimuli as function of the semantic category. *Cognitive Brain Research*, 12, 475–478.
- Viggiano, M. P., Constantini, A., Vannucci, M., & Righi, S. (2004). Hemispheric asymmetry for spatially filtered stimuli belonging to different semantic categories. *Cognitive Brain Research*, 20, 519–524.
- Zhao, L., & Bentin, S. (in press). Own- and other-race categorization of faces by race, gender, and age. *Psychonomic Bulletin and Review*.

Received September 3, 2007

Revision received July 12, 2008

Accepted July 16, 2008 ■