

Does Face Inversion Change Spatial Frequency Tuning?

Verena Willenbockel and Daniel Fiset
University of Victoria

Alan Chauvin
Université Charles-de-Gaulle Lille III

Caroline Blais and Martin Arguin
Université de Montréal

James W. Tanaka and Daniel N. Bub
University of Victoria

Frédéric Gosselin
Université de Montréal

The authors examined spatial frequency (SF) tuning of upright and inverted face identification using an SF variant of the Bubbles technique (F. Gosselin & P. G. Schyns, 2001). In Experiment 1, they validated the SF Bubbles technique in a plaid detection task. In Experiments 2a–c, the SFs used for identifying upright and inverted inner facial features were investigated. Although a clear inversion effect was present (mean accuracy was 24% higher and response times 455 ms shorter for upright faces), SF tunings were remarkably similar in both orientation conditions (mean $r = .98$; an SF band of 1.9 octaves centered at 9.8 cycles per face width for faces of about 6°). In Experiments 3a and b, the authors demonstrated that their technique is sensitive to both subtle bottom-up and top-down induced changes in SF tuning, suggesting that the null results of Experiments 2a–c are real. The most parsimonious explanation of the findings is provided by the quantitative account of the face inversion effect: The same information is used for identifying upright and inverted inner facial features, but processing has greater sensitivity with the former.

Keywords: face perception, identification, spatial frequency, inversion effect

Human adults are able to recognize thousands of faces rapidly and effortlessly, as long as the faces are seen upright (e.g., Maurer, Le Grand, & Mondloch, 2002). When stimuli are rotated by 180° in the picture plane, face recognition accuracy decreases considerably (e.g., Hochberg & Galper, 1967; Yin, 1969) and response latencies increase (e.g., Diamond & Carey, 1986). This drop in recognition performance was found to be disproportionately larger

for faces than for other mono-oriented objects (e.g., houses and airplanes) in untrained individuals and is commonly referred to as the *face inversion effect* (FIE; e.g., Boutet, Collin, & Faubert, 2003; Leder & Carbon, 2006; Robbins & McKone, 2007; Yin, 1969; reviews in Rossion & Gauthier, 2002, and Valentine, 1988). The FIE can be observed for unfamiliar and familiar faces (Collishaw & Hole, 2000; Rock, 1974; Yarmey, 1971), full faces (e.g., Boutet et al., 2003; Goffaux & Rossion, 2006, Experiment 1), or inner facial features (faces shown through an elliptical aperture; e.g., Goffaux & Rossion, 2006, Experiments 2 and 3; Robbins & McKone, 2003; Sekuler, Gaspar, Gold, & Bennett, 2004), and in a variety of experimental conditions; for instance, it has been demonstrated using blocked or randomized presentation of upright and inverted faces, in old/new recognition tasks, and in matching tasks with or without delay (for reviews, see Rossion, 2008; Rossion & Gauthier, 2002). The FIE has been described as one of the most robust phenomena in the face processing literature.

Given that upright and inverted faces have the same complexity and are almost identical in their low-level properties, such as luminance, contrast, and spatial frequencies (only phase information differs), the difficulty of recognizing upside-down faces cannot easily be attributed to stimulus properties per se. Typically, the FIE has been attributed to qualitative processing differences between upright and inverted faces, that is, inversion is thought to disrupt certain face processes more than others (e.g., Rossion, 2008). In particular, it has been suggested that inversion mainly impairs the perception of the relative distances between facial features (Diamond & Carey, 1986; Leder & Bruce, 1998, 2000;

Verena Willenbockel, Daniel Fiset, James W. Tanaka, and Daniel N. Bub, Department of Psychology, University of Victoria, BC, Canada; Alan Chauvin, Laboratoire URECA, Université Charles-de-Gaulle Lille III, Lille, France; Caroline Blais, Martin Arguin, and Frédéric Gosselin, Département de Psychologie, Université de Montréal, QC, Canada.

Portions of this research were submitted in thesis form to the Faculty of Graduate Studies at the University of Victoria. The research was supported by a graduate scholarship to Verena Willenbockel from the German Academic Exchange Service (DAAD), by a scholarship from the James S. McDonnell Foundation (Perceptual Expertise Network) and a postdoctoral scholarship from the Fonds Québécois de Recherche en Nature et Technologies (FQRNT) to Daniel Fiset, by a FQRNT graduate scholarship to Caroline Blais, by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Frédéric Gosselin and Martin Arguin and to James Tanaka, and by grants from FQRNT to Martin Arguin, Frédéric Gosselin, and Pierre Jolicoeur. We thank Bruno Rossion and Valérie Goffaux for providing the face images of Set 2.

Correspondence concerning this article should be addressed to Frédéric Gosselin or Daniel Fiset, Département de psychologie, Université de Montréal, C.P. 6128, Succ. Centre-ville, Montréal, QC H3C 3J7, Canada. E-mail: frederic.gosselin@umontreal.ca or daniel.fiset@umontreal.ca

Rhodes, 1988) or holistic processing—“the simultaneous integration of the multiple features of a face into a single perceptual representation” (Rossion, 2008, p. 275; see also Farah, Tanaka, & Drain, 1995; Sergent, 1984; Tanaka & Farah, 1993). Several other studies have demonstrated differential effects of inversion on the processing of the local facial features themselves versus their configuration or integration into a holistic representation (e.g., Bartlett & Searcy, 1993; Barton, Keenan, & Bass, 2001; Collishaw & Hole, 2000; Freire, Lee, & Symons, 2000; Leder, Candrian, Huber, & Bruce, 2001; Le Grand, Mondloch, Maurer, & Brent, 2001; Rhodes, Brake, & Atkinson, 1993; Rossion & Boremanse, 2008; Tanaka & Sengco, 1997; Thompson, 1980; Young, Hellawell, & Hay, 1987), thus supporting the qualitative view of the FIE.

Despite a wide acceptance of the qualitative account, it has also been proposed that the FIE arises from quantitative processing differences. Using a response classification method (Ahumada & Lovell, 1971; Eckstein & Ahumada, 2002), Sekuler et al. (2004) found that the eye and eyebrow regions contain the features correlated linearly with the discrimination of both upright and inverted faces, and that the correlations associated with upright faces were greater than those associated with inverted faces. These findings are in line with previous behavioral studies that identified the eye and eyebrow regions as most important for face identification (e.g., Davies, Ellis, & Shepherd, 1977; Gold, Sekuler, & Bennett, 2004; Gosselin & Schyns, 2001, 2005; Haig, 1985, 1986; Schyns, Bonnar, & Gosselin, 2002; Shepherd, Davies, & Ellis, 1981; Walker-Smith, Gale, & Findlay, 1977). They are also consistent with eye-tracking data demonstrating that gaze is especially attracted by the eyes of both upright and inverted faces (e.g., Williams & Henderson, 2007; but see Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006, for a demonstration of more random global scan patterns and more fixations on lower face parts in the inverted condition). Furthermore, Sekuler et al. reported that contributions of nonlinear processes to performance were small and similar for both face orientations. The authors concluded from their classification image results that a quantitative, not a qualitative, difference underlies the FIE (see also Gaspar, Bennett, & Sekuler, 2008; Riesenhuber, Jarudi, Gilad, & Sinha, 2004; Valentine, 1988; Yovel & Kanwisher, 2004; but see Rossion, 2008)—specifically, that the same facial information is processed with less sensitivity when faces are upside-down.

The classification images of Sekuler et al. (2004), however, are insufficient to determine whether the same cues subtend upright and inverted face recognition beyond the spatial domain. For example, they cannot rule out that the eye and eyebrow regions revealed in the upright and inverted classification images hide different patterns of spatial frequency (SF) use (e.g., 8 cycles per face width for upright eye and eyebrow regions and 16 cycles per face width for the same regions but inverted). It thus remains possible that qualitative processing differences can be found in the SF domain.

Spatial Frequency Tuning for Face Identification

The human visual system analyzes the complex luminance variations that make up the visual stimulus with discrete channels, each tuned to a specific SF range (see De Valois & De Valois, 1990, for a review). High SFs represent the fine-grained informa-

tion in a stimulus, such as the eyelashes or edges of the mouth, and low SFs convey coarse information, such as luminance blobs and blurred shapes (see Morrison & Schyns, 2001, and Ruiz-Soler & Beltran, 2006, for reviews). Several studies found that a narrow band of intermediate SFs centered between 7 and 16 cycles per face (cpf) width is particularly important for upright face recognition (e.g., Costen, Parker, & Craw, 1994, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999).

For the SF tuning of inverted face identification, mixed findings have been obtained. Collishaw and Hole (2000) demonstrated that blurred faces (i.e., faces containing only low-SF information) could still be recognized above chance level, unless they were presented upside-down. Because inversion of a scrambled face (a condition of comparable difficulty) did not decrease performance below chance, the findings were attributed to a disruption of the processing of the relative distances between facial features (by inversion) and the featural information (by blurring) rather than task difficulty. These findings suggest that the FIE might be particularly large for low-pass filtered faces. Although their study did not directly pertain to face inversion, Goffaux and Rossion (2006, Experiments 2 and 3) demonstrated in a composite face paradigm that inversion costs in both accuracy and RT were larger for low-SF faces (<8 cpf) compared with broad-spectrum and high-SF (>32 cpf) stimuli. Moreover, Nakayama (2003) reported that face discrimination (in a four-choice task) was most degraded by noise of approximately 12 cpf, with a much broader masking function for inverted than for upright faces. Finally, using an old/new recognition task, Boutet et al. (2003, Experiment 1) found FIEs of comparable magnitude on accuracy for broadband faces, band-pass filtered faces in a range of relatively low SFs (1.25–5 cpf), and intermediate-SF (5–20 cpf) faces, but not for high-SF (20–80 cpf) faces, which did not lead to a significant FIE. These findings, albeit somewhat inconsistent, support the qualitative view of the FIE.

In contrast to the above, Boutet et al. (2003, Experiments 2A and B) found a comparable FIE on accuracy in all SF band-pass filter conditions (1.25–5 cpf, 5–20 cpf, 20–80 cpf) in a sequential matching paradigm and concluded—considering the results of their Experiments 1–3—that SF filtering had little impact on the FIE. Furthermore, Collin, Liu, Troje, McMullen, and Chaudhuri (2004) demonstrated that varying the degree of SF overlap between two faces influences accuracy in a matching task similarly for upright and inverted faces. Finally, using critical band masking—a technique that measures signal thresholds for stimuli to which high-pass (or low-pass) filtered white Gaussian noise is added at different cutoffs—Gaspar, Sekuler, and Bennett (2008) showed that the same narrow SF band (approximately 1.5 octaves wide and centered at about 7 cpf) was used in two 10-choice identification tasks with both upright and inverted faces. These results are consistent with a quantitative view of the FIE. In sum, both orientation dependency and independency of SF use has been reported in the literature.

Here, we re-examined whether the identification of upright and inverted faces is mediated by different SFs using a novel SF variant of the Bubbles technique (Fiset, Blais, Gosselin, & Schyns, 2006; Gosselin & Schyns, 2001; see McCotter, Gosselin, Sowden, & Schyns, 2005, for a distinct attempt at applying Bubbles to SFs). The Bubbles technique has been applied to full-spectrum images (e.g., Gosselin & Schyns, 2001, Experiment 1) as well as band-

pass filtered images (e.g., Gosselin & Schyns, 2001, Experiment 2; Schyns et al., 2002). However, it has not previously been used to sample the SF content of stimuli. We randomly varied the availability of SFs on a trial-by-trial basis, thereby testing the contribution of each SF to performance independently. After a sufficient number of trials, we performed multiple linear regressions on the random SF filters and response accuracy or response time (RT).

The SF Bubbles technique allowed us to derive the precise SF tuning curves for upright and inverted face identification. In comparison, traditional high-pass, low-pass, or band-pass filtering techniques (e.g., Boutet et al., 2003; Goffaux & Rossion, 2006) provide only crude estimates of SF tuning. The main strength of the SF Bubbles technique in the present context is that it minimizes the risk that participants adapt to a predictable stimulus manipulation (e.g., low-, band-, or high-pass filtering or critical band masking; see Sowden & Schyns, 2006, for evidence of “channel surfing”) by randomly sampling multiple SFs simultaneously on a trial-by-trial basis.

The present research comprised six experiments. The first experiment was designed to assess the validity of the SF Bubbles technique. Using a plaid (i.e., the sum of two sine wave gratings) detection task, we verified that the SF Bubbles method can uncover precisely the diagnostic SFs. The next three experiments, which constitute the main experiments, examined which SFs are used for the accurate (Experiments 2a and c) and fast (Experiment 2b) identification of upright and inverted faces. The experiments employed two sets of 20 grayscale face photos (10 identities \times 2 exemplars per set) that were cropped to an elliptical shape. We decided to focus on inner facial features to bridge the qualitative and quantitative accounts of the FIE as much as possible (e.g., Gaspar, Sekuler, & Bennett, 2008; Goffaux & Rossion, 2006, Experiments 2 and 3; Robbins & McKone, 2003; Sekuler et al., 2004, used faces revealed through an elliptical aperture). To anticipate our main result, we find no difference in SF use between the upright and inverted condition. To rule out that this null result was due to an insensitivity of the SF Bubbles technique to bottom-up or top-down influences on SF tuning, we carried out two more experiments. Experiment 3a re-examined SF tuning in the identification task of Experiment 2a as a function of stimulus size, which is known to influence SF tuning in a bottom-up fashion (e.g., Majaj, Pelli, Kurshan, & Palomares, 2002). In Experiment 3b, task demands were modified (gender or happy/neutral discrimination) to modulate SF tuning in a top-down fashion (e.g., Schyns & Oliva, 1999). In both cases, we revealed subtle differences in SF use, confirming that the SF Bubbles technique is sensitive to bottom-up- and top-down-induced changes in SF tuning.

Experiment 1

The purpose of the first experiment was to determine whether the SF Bubbles technique can reveal precisely the SFs that convey the information that is diagnostic for the task. We employed a plaid (i.e., the sum of two SFs) detection task: If the SF Bubbles technique works adequately, then we should be able to recover the SFs in the plaid. This experiment is also meant as an illustration of the SF Bubbles technique.

Method

Participants. One male and two female University of Victoria students (22–25 years old; $M = 24$ years) participated in Experiment 1. All participants had normal or corrected-to-normal vision, and two were naïve to the purpose of the experiment. The third participant was the first author of the article (Observer 1). Participants gave informed consent approved by the University of Victoria Human Research Ethics Committee.

Apparatus. Experiments 1, 2a, 2b, 3a, and 3b were run on a dual core 2.93 GHz PC at the University of Victoria. Stimuli were displayed on a 22-in. Viewsonic CRT monitor that was calibrated to allow a linear manipulation of luminance. The resulting corrected table contained 154 luminance levels, ranging from 0.3 cd/m^2 to 98.7 cd/m^2 . The background luminance was 49.3 cd/m^2 . The monitor refresh rate was 85 Hz for all experiments, and the resolution was set to 1,024 \times 768 pixels (except in Experiment 3a). All experiments were programmed in MATLAB (Natick, MA) using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Procedure. Participants were instructed to perform a plaid detection task. The original plaid comprised a horizontal sine wave grating with an SF of 10 cycles per image (cpi) and a vertical sine wave grating with an SF of 45 cpi (Figure 1) and had a size of 256 \times 256 pixels. On “signal present” trials (50% of trials), the SFs of the plaid were randomly sampled (see next section) and displayed embedded in white Gaussian noise. On “signal absent” trials, a white Gaussian noise field of 256 \times 256 pixels was displayed. Each trial began with a central fixation cross lasting 435 ms, followed by the stimulus presented for 870 ms, and then by a homogeneous mid-gray field that remained on the screen until the observer responded by pressing an appropriate key on a computer keyboard. Signal present and signal absent trials occurred in random order. No feedback was provided. Each observer performed eleven 100-trial blocks with breaks between blocks. Participants were seated in a dark room and a chin rest was used to maintain viewing distance at 53 cm; stimuli subtended a visual angle of $10.2^\circ \times 10.2^\circ$.

Spatial frequency bubbles. All experiments reported in this article revealed SF use by employing the SF Bubbles technique. This section serves two purposes: (a) It describes the SF Bubbles technique in general, and (b) it illustrates the use of the technique in Experiment 1.

On each trial, the SF information of a stimulus was sampled randomly as illustrated in Figure 1. First, the square base stimulus was padded to minimize edge artifacts in the SF domain. It was centered on a uniform gray field of the stimulus’ background luminance and twice its size. In Experiment 1, for example, the plaid of size 256 \times 256 pixels was padded with a midluminance background of size 512 \times 512 pixels. Second, the padded stimulus was Fourier transformed using functions from the Image Processing Toolbox for MATLAB. The quadrants of the Fourier image were shifted so that low SFs occupied the central region of the complex (i.e., real + imaginary number) amplitude matrix. Third, a random filter was constructed. The construction of this filter involved the following steps: (a) A binary random vector of $2wk$ elements was created, where w was the stimulus width and k a constant that determined the smoothness of the sampling (the

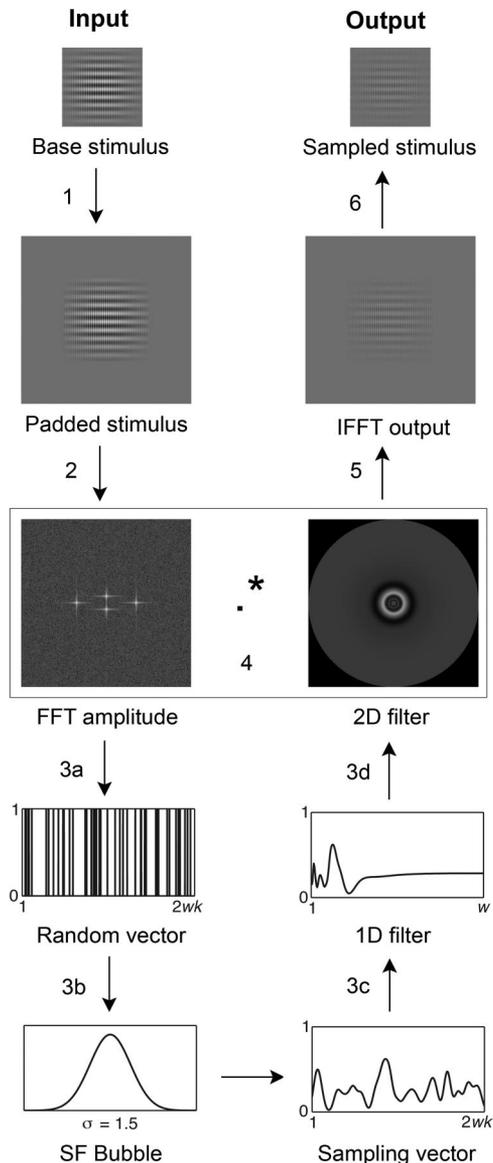


Figure 1. Illustration of the spatial frequency (SF) Bubbles technique. (1) Padding of the base stimulus with a uniform gray background. (2) Fast Fourier transform (FFT) of the padded stimulus. (3) Construction of a random SF filter: (a) creation of a binary random vector of length $2wk$ (w = image width; $k = 20$) consisting of 45 ones among zeros; (b) convolution of the random vector with a Gaussian kernel (an “SF bubble”), resulting in a random sampling vector; (c) log-scaling of sampling vector (see text for details), resulting in a one-dimensional (1D) filter; (d) construction of a 2D filter by rotating the 1D filter about its origin. (4) SF filtering by dot-multiplying the 2D filter with the padded stimulus’ complex FFT amplitudes. (5) Inverse fast Fourier transform (IFFT) and (6) cropping of the output to create the SF sampled experimental stimulus.

higher k , the smoother); k was arbitrarily set to 20 for all the experiments reported in this article. In Experiment 1, the random vector thus had 10,240 elements ($2 \times 256 \times 20$). The vector contained zeros among b ones that were randomly distributed (with repetition); b determined the number of SF bubbles (see

below) and was arbitrarily set to 45. (b) To create a smooth filter, the binary vector was convolved with a Gaussian kernel, referred to as an *SF bubble*. The standard deviation of the SF bubble was arbitrarily set to 1.5 and the maximum to 0.125; all values of the resulting vector above 1 were reduced to 1. The convolution resulted in a “sampling vector” consisting of b randomly located SF bubbles. (c) To ensure that the sampling vector approximately fit the SF sensitivity of the human visual system (see De Valois & De Valois, 1990, for a review), we subjected the smoothed vector to a logarithmic transformation: w elements of the vector were sampled according to the function of

$$f(x) = e^{(x-1)\frac{\ln(kw)}{(w-1)}} + a,$$

with $x = [1:w]$ and $a = kw/2$. The constant a was arbitrarily chosen and prevented low and high SFs being sampled less often than intermediate SFs. (d) The resulting w -element filter was then rotated about its origin to create an isotropic random two-dimensional filter of size $w \times w$.

Fourth, filtering was carried out by dot-multiplying the two-dimensional filter with the complex amplitude of the padded base stimulus before subjecting the result to the inverse Fourier transform. We constructed the experimental stimuli by cropping the central $w \times w$ pixel region of the filtered image. White Gaussian noise was added to the SF sampled stimulus to adjust performance. The $w \times w$ noise field was multiplied by $1 - c$, with c ranging from 0 to 1 and added to the image multiplied by c . The value of c was increased or decreased on a block-by-block basis by the experimenter (in increments of .02; Experiments 2a–c and 3a) or on a trial-by-trial basis using QUEST (Watson & Pelli, 1983; Experiments 1 and 3b). For example, in Experiment 1, performance was maintained at 75% correct.

To find out which SFs drove the observers’ correct and incorrect responses or response times, we performed a multiple linear regression on the random binary vector (see above) and transformations of the observers’ correct and incorrect responses or RTs. Here, a multiple linear regression is linearly related to summing all sampling vectors weighted by the transformed responses. Throughout this article, correct and incorrect responses were transformed as follows: Correct responses were given a value of $1 - P(\text{correct})$ and incorrect responses a value of $-P(\text{correct})$. In Experiment 1, for example, correct responses were assigned a value of 0.25 and incorrect responses a value of -0.75 . Similarly, fast responses (RTs shorter than the median RT of the corresponding block) were given a value of 1 and slow responses (those longer than the median RT) were given a value of -1 .

The vector of w regression coefficients—referred to as a *classification vector*—was then transformed into Z scores for each observer. A group classification vector can be computed by summing the classification vectors of all observers and by dividing the resulting vector by \sqrt{n} , with n equal to the number of observers. A pixel test was used to determine a statistical threshold (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). Note that because of padding, the x th element of the classification vector corresponds to $x/2$ cycles per base stimulus width; in this article, all SFs are given either relative to the base stimuli (Experiment 1) or to face width (Experiments 2a–c and 3a and b).

Results and Discussion

The first block was considered as practice block and was excluded from the analysis. Results are shown in Figure 2, which plots the Z scores across all SFs (classification vectors) as well as the significance threshold ($p < .05$, $S_r = 256$, $FWHM = 3.53$, $Z_{crit} = 3.45$; for details, see Chauvin et al., 2005). Individual and group results were very similar ($r_s = .93$, $.99$, and $.96$); we thus report the group results only. Two significant peaks occurred: the first peak at 10 cpi ($Z_{max} = 10.50$) and the second at 45 cpi ($Z_{max} = 6.87$), with an octave width of 0.42 and 1.39, respectively. The peaks correspond to the two SFs of the plaid, thus demonstrating that the SF Bubbles technique can accurately reveal the SF information that drives the observers' responses.

Experiments 2a–c

Experiments 2a–c were designed to investigate which SFs underlie the identification of upright and inverted faces. For upright faces, several studies demonstrated that a narrow band of intermediate SFs (i.e., centered between 7 and 16 cpf, depending on the size of the face stimuli) is particularly important for recognition (e.g., Costen et al., 1994, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999). For inverted face identification, less is known about the exact SF range used. Previous results indicate that face inversion might be particularly detrimental for low-SF faces (e.g., Collishaw & Hole, 2000; Goffaux & Rossion, 2006), that the SF band used for inverted faces might be much broader than for upright faces (Nakayama, 2003), or that SF tuning might be very similar for upright and inverted face identification (Gaspar, Sekuler, & Bennett, 2008). We tested the independent contribution of each SF to the identification of upright and inverted inner facial features, allowing us to reconstruct the precise SF filters used for effective identification. We used both accuracy (Experiments 2a and c) and RT (Experiment 2b) as measures of effective identification and employed the SF Bubbles technique described above to reveal the SF tuning curves.

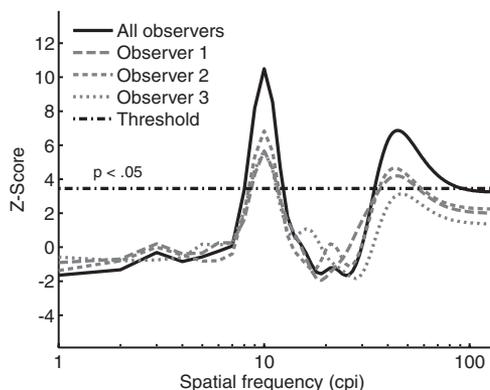


Figure 2. Individual and group classification vectors obtained in the plaid detection task (Experiment 1). The spatial frequency (SF) Bubbles technique revealed significant peaks at 10 cycles per image (cpi) and at 45 cpi ($Z_{crit} = 3.45$, $p < .05$), thus accurately showing the diagnostic SFs (i.e., the SFs of the plaid).

Method

Participants. In total, 15 university students (10 women) ages 19–35 years ($M = 25.5$ years) took part in Experiments 2a–c. Four participants completed both Experiments 2a and b, and 1 of them additionally took part in Experiment 2c. One student participated in Experiment 2a only and 1 in 2b only. Nine participants took part in Experiment 2c only. All observers had normal or corrected-to-normal vision. Two participants are among the authors of this article, and 13 participants were naïve to the purpose of the study. They received course credit or were paid as compensation.

Apparatus. For Experiments 2a and b, see the corresponding section of Experiment 1. Experiment 2c was run on a dual core 2.2 GHz PC at the Université de Montréal. Stimuli were presented on a 22-in. Hewlett Packard CRT monitor that was calibrated to allow a linear manipulation of luminance. The corrected table contained 117 luminance levels, ranging from 0.3 cd/m^2 to 70.7 cd/m^2 ; the background luminance was 29.3 cd/m^2 . The monitor refresh rate was 85 Hz, and the resolution was set to $1,024 \times 768$ pixels. The experiment was run in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Stimuli. Two sets of grayscale face photos of 256×256 pixels served as base stimuli. Each set comprised two exemplars of 10 faces to make it less likely that observers followed a template-matching strategy. In Set 1 (Experiments 2a and b), the exemplars showed two different expressions (neutral and happy). In Set 2 (Experiment 2c)—a subset of the base faces used by Goffaux and Rossion (2006)—both exemplars had a neutral expression, but the faces were shown from slightly different angles. The main inner facial features (eyes, eyebrows, nose, and mouth) were aligned within each stimulus set using rotation, translation, and scaling. It is important to note that these affine transformations did not alter the shape of facial features or the relative distance between them. The faces were presented behind a midluminance homogeneous field through an elliptical aperture with a horizontal diameter of 158 or 121 pixels and a vertical diameter of 239 or 175 pixels for the first and second stimulus set, respectively (see Figures 3a and 3b). Thereby, only the inner facial features were revealed, and the minor and major axes of the ellipse contained no information relevant to the task at hand. This mode of stimulus presentation was also chosen by Gaspar, Sekuler, and Bennett (2008), Robbins and McKone (2003), and Sekuler et al. (2004). In contrast, Goffaux and Rossion (2006, Experiments 2–4) used different elliptical apertures to reveal the inner facial features of their stimuli. Faces within each stimulus set were equated in mean luminance, contrast, and energy at each SF. Face width subtended a visual angle of 6.5° (Set 1) or 5.7° (Set 2). The base stimuli were presented upright or inverted (rotated 180° in the image plane).

Practice phase. Participants learned to associate the faces with common names (e.g., Mary, John, Peter) from printed grayscale pictures displayed along with names. When the participants were confident that they could identify all faces, the practice session began. Participants were seated in a dark room and a chin rest maintained them at a 53-cm viewing distance from the screen in Experiments 2a and b and at a 45-cm viewing distance for the smaller faces in Experiment 2c. Upright and inverted base stimuli were presented in separate 100-trial blocks, starting with upright faces. Each trial began with the appearance of a central fixation cross on the screen for 435 ms, followed by an upright or an

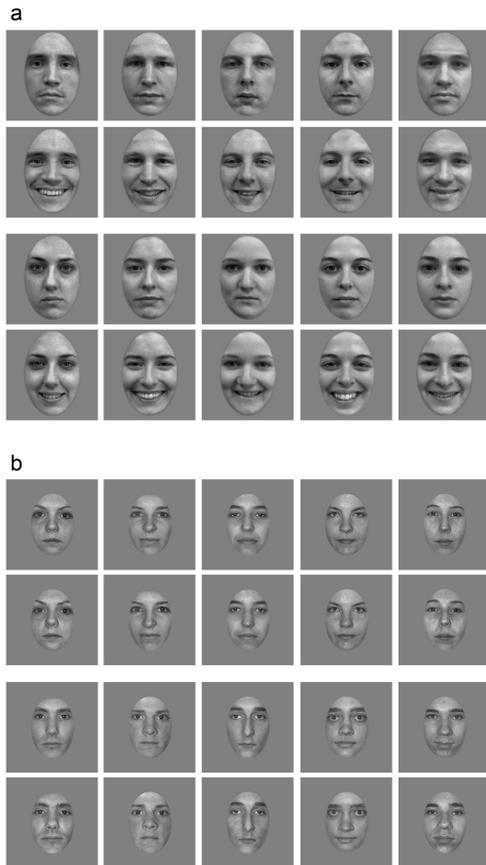


Figure 3. (a) Base stimuli used in Experiments 2a, 2b, and 3a. Faces of 10 identities \times 2 expressions (neutral, happy) were shown. (b) Base stimuli used in Experiment 2c (from the base set of Goffaux & Rossion, 2006). Faces displayed 10 neutral identities \times 2 slightly different viewpoints.

inverted face presented for 435 ms, and then by a homogeneous mid-gray field that remained on the screen until the observer responded with a key press. Each of the keys (numerals 0 to 9) on a regular computer keyboard was associated with a particular face name. When participants responded incorrectly, auditory feedback was provided (a brief 3000-Hz pure tone). The first part of the practice session was completed when accuracy for upright faces was above 95% correct for two successive blocks of 100 trials; the second part was completed when the same criterion was reached for inverted faces. On average, participants needed 6.40 or 6.00 practice blocks in the upright condition and 17.00 or 10.50 blocks in the inverted condition with the first or second stimulus set, respectively. Finally, participants performed six additional practice blocks (three with upright and three with inverted faces) in which white Gaussian noise was added to the full-SF spectrum faces in preparation for the experimental blocks.

Experimental phase. In the experimental phase of Experiment 2a, each participant was presented with a total of 2,100 upright and 2,100 inverted face stimuli. Upright and inverted faces were presented in separate 100-trial blocks, starting with upright stimuli and then alternating with inverted stimuli. Accuracy was measured in the same 10-choice identification task as during the practice. The experimental trials differed from the practice trials on

the following: (a) The SFs of the base stimuli were randomly sampled (Figure 4; for details, see the *Spatial frequency bubbles* section), (b) no feedback was given, and (c) performance in the upright blocks was maintained between 75% and 85% correct by adjusting the quantity of additive noise block per block. The same amount of noise was used in the following inverted blocks. We chose to equate the quantity of additive noise across conditions instead of accuracy (the latter is what Sekuler et al., 2004, and Gaspar, Sekuler, & Bennett, 2008, did) because in “real life” upright and inverted faces contain the same amount of information.

The RT version (Experiment 2b) followed Experiment 2a and differed from it only in four respects: (a) Face stimuli remained on the screen until a response was made; (b) participants named aloud the identities of the faces, and a voice key was used to measure response latency; (c) after each trial, the experimenter typed the participants’ response using a computer keyboard; and (d) performance was maintained above 90% correct in the upright condition block per block.

Experiment 2c served as control experiment and differed from Experiment 2a only in three respects: (a) The second stimulus set was used (Figure 3b), (b) the number of trials per observer was reduced to 400 per condition, and (c) the sigma of the Gaussian

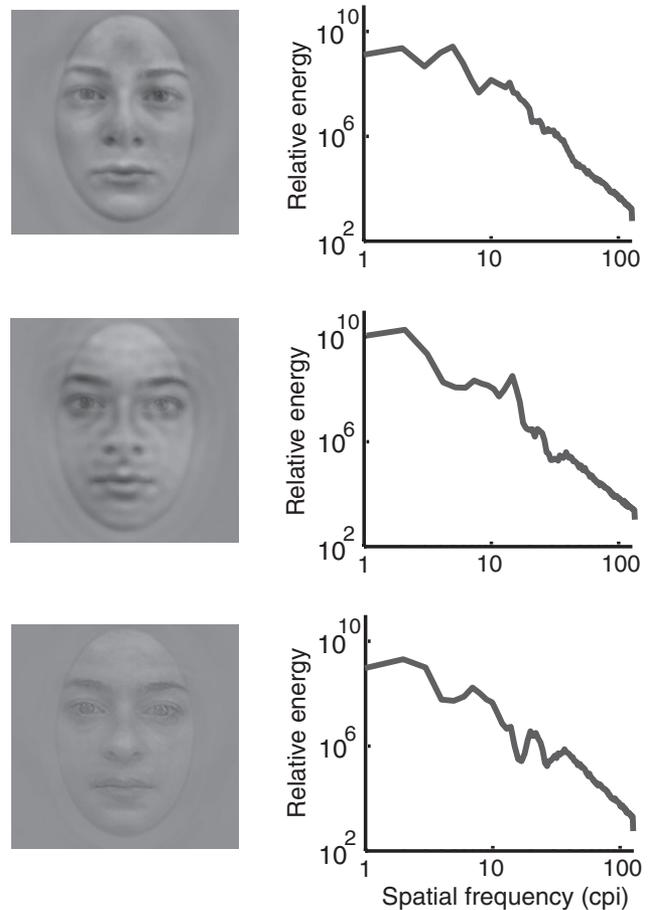


Figure 4. Three sample stimuli after filtering with the spatial frequency (SF) Bubbles technique. Fourier energy averaged across orientations is plotted as a function of SF in cycles per image (cpi).

filter was increased to 1.8 octaves. The goal of this experiment was to see whether the results obtained with the first set of stimuli would generalize to another set of faces. Furthermore, we wanted to see whether the results could be replicated with a relatively small number of trials per observer.

Results and Discussion

In Experiments 2a–c, the first 100-trial block in each orientation condition was considered as practice and was therefore excluded from the analyses. In Experiment 2a, accuracy across participants was significantly higher for upright ($M = 80.52\%$, $SD = 1.35$) than for inverted faces ($M = 52.43\%$, $SD = 6.32$), $t(4) = 8.56$, $p < .01$. In Experiment 2b, where accuracy was adjusted to above 90% for upright faces ($M = 92.24\%$, $SD = 2.23$), accuracy for inverted faces was again lower ($M = 72.12\%$, $SD = 6.12$), $t(4) = 6.81$, $p < .01$. Furthermore, RTs on correct trials were significantly shorter in the upright ($M = 1,479.83$ ms, $SD = 315.29$) than in the inverted condition ($M = 1,935.01$ ms, $SD = 410.95$), $t(4) = -6.90$, $p < .01$. In Experiment 2c, accuracy was also significantly higher for upright ($M = 79.37\%$, $SD = 5.81$) than for inverted faces ($M = 55.43\%$, $SD = 7.51$), $t(9) = 14.87$, $p < .01$. In sum, all versions of the experiment exhibited a clear FIE. Figure 5 shows the mean accuracy (Experiments 2a–c) and RTs (Experiment 2b) for each block.

To reveal the SF ranges that led to accurate (Experiments 2a and c) and fast (Experiment 2b) face identification in the upright and the inverted conditions, we performed multiple linear regressions on the sampling vectors per orientation condition per observer and on the appropriate regressor (see the *Spatial frequency bubbles* section). Because individual and group classification vectors were very similar in Experiments 2a and 2b (average correlations including both orientation conditions of $r = .86$ in Experiment 2a and $r = .77$ in Experiment 2b), and because Experiment 2c was based on a relatively small number of trials per observer, we report group results only. The group classification vectors for the upright and inverted conditions and their Z -transformed difference are shown in Figures 6, 7, and 8. The upright group classification vector in Experiment 2a showed a significant SF band of 2.00 octaves and dual peaks, one at 7.14 cpf ($Z_{\max} = 8.20$, $p < .05$, $S_r = 256$, $FWHM = 3.53$, $Z_{\text{crit}} = 3.45$) and the other at 12.14 cpf ($Z_{\max} = 8.10$). Similarly, in the inverted condition, a 2.00 octaves wide SF band peaking at 7.14 cpf ($Z_{\max} = 8.68$) and 11.07 cpf ($Z_{\max} = 7.72$) was significant. In Experiment 2b, the group classification vector for the upright condition revealed a significant SF range of 1.94 octaves peaking at 8.57 cpf ($Z_{\max} = 6.97$) and 12.86 cpf ($Z_{\max} = 5.77$). In the inverted condition, an SF range of 1.30 octaves peaking at 12.14 cpf ($Z_{\max} = 5.06$) was significant. In Experiment 2c, the upright classification vector reached significance for an SF band of 1.00 octave with a maximum Z score at 7.53 cpf ($Z_{\max} = 4.34$, $p < .05$, $S_r = 256$, $FWHM = 4.24$, $Z_{\text{crit}} = 3.40$). The classification vector for the inverted condition was significant for an SF band of 1.06 octaves, with a peak at 8.06 cpf ($Z_{\max} = 5.06$, $p < .05$, $S_r = 256$, $FWHM = 4.24$, $Z_{\text{crit}} = 3.40$). None of the difference classification vectors reached statistical significance.

The correlations between classification vectors obtained on accuracy and RTs using the same stimuli (Experiments 2a and b) were very high in both the upright conditions ($r = .96$) and the

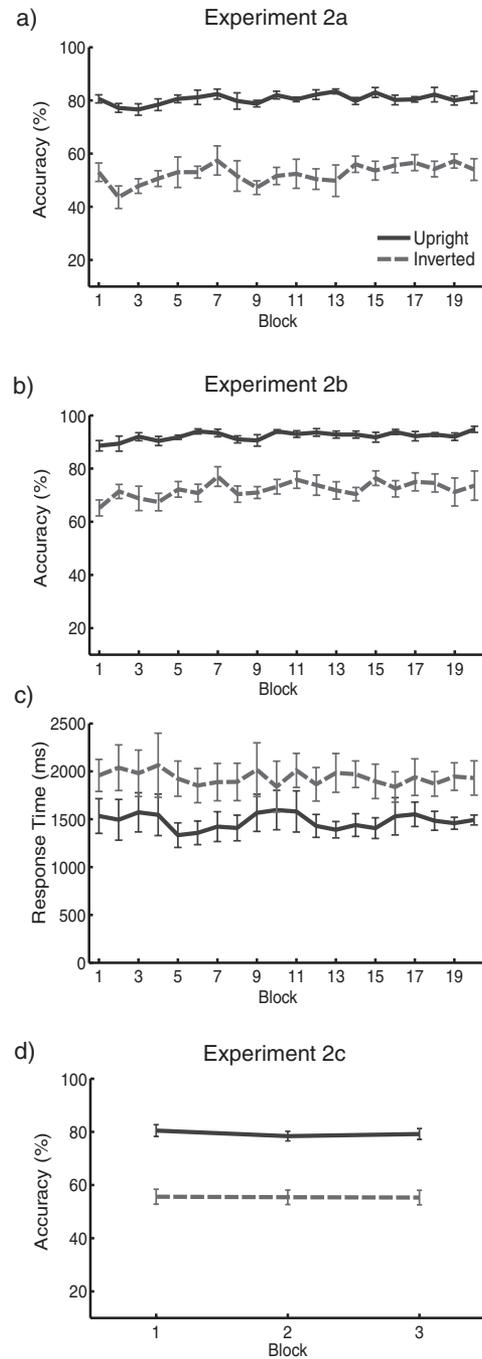


Figure 5. Mean accuracy across observers over the 20 blocks of Experiments 2a (a) and 2b (b), as well as over the three blocks of Experiment 2c (d). Mean response times (RTs) on correct trials over the 20 blocks of Experiment 2b are shown in (c). Error bars give the standard errors. Both accuracy and RT show a clear face inversion effect.

inverted conditions ($r = .93$). For the experiments measuring accuracy but using different stimulus sets (Experiments 2a and c), the correlations were high as well ($r = .84$ for upright faces and $r = .87$ for inverted faces). Similarly, we found high correlations between Experiments 2b and c (based on different stimulus sets

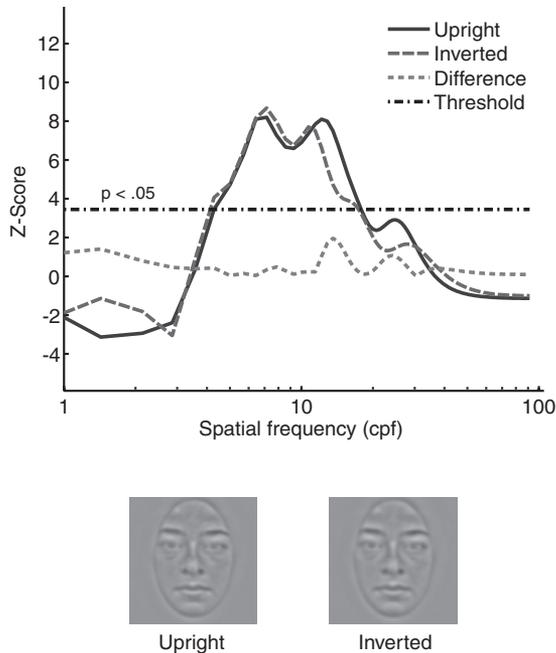


Figure 6. Accuracy group classification vectors obtained for upright and inverted face identification in Experiment 2a. The two faces at the bottom show the spatial frequencies (cpf = cycles per face) that reached statistical significance ($Z_{crit} = 3.45$, $p < .05$) in both orientation conditions. There was no significant difference between the two classification vectors.

and different response measures) with $r = .87$ for upright and $r = .95$ for inverted faces. Most important, the correlation between the classification vectors for upright and inverted faces was very high in Experiments 2a ($r = .97$), 2b ($r = .95$), and 2c ($r = .85$), and the correlation between the average of the upright and inverted classification vectors of Experiment 2a–c was even higher ($r = .98$), strongly suggesting that the same SF band was used for identifying upright and inverted faces. To best estimate the center and width of this SF band, we summed all classification vectors for Experiments 2a–c and fitted a Gaussian density function to the logarithm of the average classification vector. The mean of the best fit—9.8 cpf—is our estimate of the center of the SF channel, and its *FWHM*—1.9 octaves—is our estimate of the width of the SF channel.

The reliance on this 1.9 octaves wide SF band centered on average on 9.8 cpf appears robust to changes in the measure employed to assess the FIE (accuracy or response time of correct trials), to changes in the face sets, and to changes in procedures (many trials and few subjects in Experiments 2a and 2b, or few trials and relatively many subjects in Experiment 2c). The SF range revealed in the present study is in accordance with the intermediate SFs identified in previous studies for upright faces (e.g., Costen et al., 1994, 1996; Gold et al., 1999; Näsänen, 1999). Our results are also consistent with the findings of Gaspar, Sekuler, and Bennett (2008) for both upright and inverted faces, and suggest that the processing difference does not lie in the extraction of cues at different SFs but rather in how the information within the same SF band is used.

Experiments 3a and b

The results of Experiments 2a–c reveal no difference in SF tuning between the upright and inverted conditions. However, we cannot rule out the possibility of this “null result” being due to a relative insensitivity of the SF Bubbles technique to reveal subtle differences in SF tuning resulting from bottom-up or top-down alterations in the visual strategies employed by observers with upright versus inverted faces (e.g., holistic vs. featural processing). We conceived two experiments to test whether the SF Bubbles technique is sensitive enough to reveal subtle differences in SF tuning in situations known to differ solely in bottom-up (Experiment 3a) or top-down (Experiment 3b) SF tuning.

It has been shown that decreasing the size of letters (Chung, Legge, & Tjan, 2002; Majaj et al., 2002) and faces (Loftus & Harley, 2005; Näsänen, 1999) induces a shift in the use of SFs toward lower SFs in a mandatory bottom-up fashion. Experiment 3a examined whether we can reveal such a bottom-up-induced SF tuning change with the SF Bubbles technique. Observers were asked to identify the same faces as in Experiments 2a and b with large versus small stimuli.

Experiment 3b was designed to assess the capacity of the SF Bubbles technique to reveal differences in SF tuning that are due solely to changes in the top-down influence of task demands. More specifically, the SF Bubbles technique was applied to two different tasks previously shown to induce different SF usage patterns (happy/neutral vs. gender discriminations; e.g., Gosselin & Schyns, 2001; Schyns & Oliva, 1999).

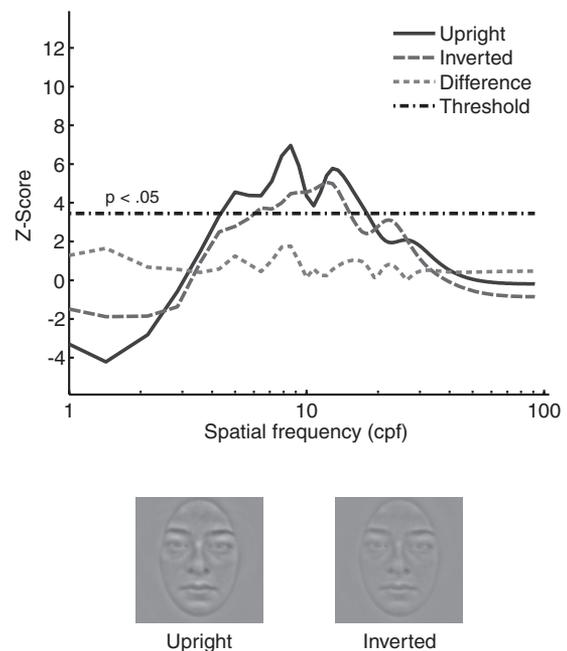


Figure 7. Response time group classification vectors obtained for upright and inverted face identification in Experiment 2b. The two faces at the bottom show the spatial frequencies (cpf = cycles per face) that reached statistical significance ($Z_{crit} = 3.45$, $p < .05$) in both orientation conditions. There was no significant difference between the two classification vectors.

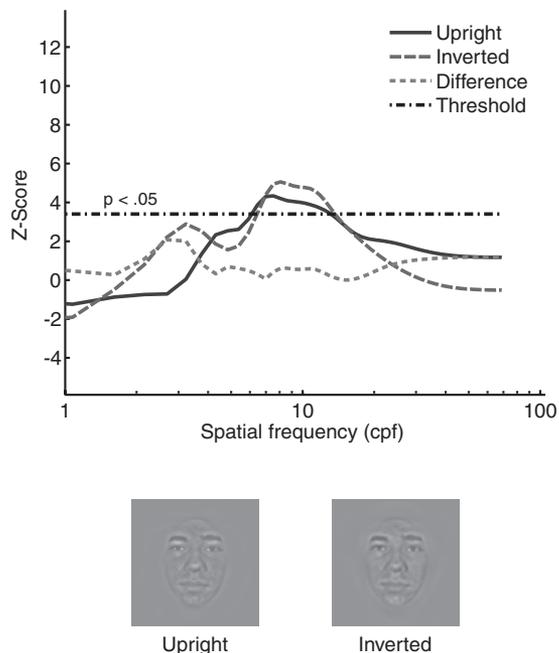


Figure 8. Accuracy group classification vectors obtained for upright and inverted face identification in Experiment 2c with stimuli from Goffaux and Rossion (2006). The two faces at the bottom show the spatial frequencies (cpf = cycles per face) that reached statistical significance ($Z_{crit} = 3.40$, $p < .05$) in both orientation conditions. There was no significant difference between the two classification vectors.

Method

Participants. For Experiment 3a, 1 male and 2 female University of Victoria students (ages between 22 and 26 years; $M = 23.7$ years) were recruited. For Experiment 3b, we recruited 40 students (31 women) between 18 and 42 years of age ($M = 19.8$ years). All observers had normal or corrected-to-normal vision. Participants were naïve to the purpose of the study, and they received course credits or were paid for participating.

Apparatus. See corresponding section of Experiment 1; for the screen settings in Experiment 3a, see next section.

Stimuli. In Experiment 3a, the base stimuli of Experiments 2a and b were used but their resolution was decreased to 128×128 pixels. In the small condition, observers saw the upright face stimuli at a screen resolution of $2,048 \times 1,536$ pixels and at a distance of 180 cm, resulting in a face width of 0.5° of visual angle. In the big condition, they saw them at a screen resolution of 640×480 pixels and at a viewing distance of 45 cm, resulting in a face width of 5.9° of visual angle.

In Experiment 3b, the same 10 neutral faces (5 males, 5 females) were used as in Experiments 2a, 2b, and 3a. The corresponding 10 happy faces, however, differed from the set previously used in that no teeth were visible (Figure 9). We chose this set to make task difficulty between gender and happy/neutral discriminations more similar; a pilot study showed ceiling effects for happy/neutral discrimination when accuracy for gender was in the 65%–75% range. Base stimuli had a resolution of 256×256 pixels and face width subtended a visual angle of 6.5° . They were normalized for

a number of low-level visual features and for the main facial feature position as in Experiments 2a, 2b, and 3a.

Procedure. In Experiment 3a, participants performed the same 10-alternative face identification task as in Experiment 2a. Each participant completed 3,300 trials per condition. We adjusted the quantity of additive white Gaussian noise (as described in the *Method* section for Experiment 1) on a block-by-block basis so that performance was approximately 80% correct in the big condition (i.e., the easiest condition). The 100-trial blocks succeeded each other as in Experiment 2a, but this time alternating between big and small rather than between upright and inverted.

Experiment 3b was divided in two parts: Each participant completed six consecutive 100-trial blocks of happy/neutral discrimination and six consecutive 100-trial blocks of gender discrimination. The first 20 participants initially completed the happy/neutral discrimination, followed by the gender discrimination; the last 20 participants completed the tasks in the opposite order. Each trial began with a central fixation cross presented for 412 ms, followed by an upright face presented for 412 ms, and then by a homogeneous mid-gray field that remained on the screen until the observer responded by pressing the appropriate key on a computer keyboard. Keys were counterbalanced across participants. For the initial task, performance was adjusted on a trial-by-trial basis by manipulating the quantity of additive noise using QUEST (Watson & Pelli, 1983). In the second task of Experiment 3b, the same experimental stimuli were used (i.e., same base faces, same sampling vector, and same amount of white Gaussian noise) as in the corresponding trials of the first task of Experiment 3b.

Results and Discussion

For each participant, the first 100-trial block of each task was considered as practice and was discarded from data analysis. The analyses were thus performed on 19,200 trials (3,200 trials per size condition \times 2 size conditions \times 3 participants) and 40,000 trials



Figure 9. Base stimuli used in Experiment 3b. Faces displayed 10 identities \times 2 expressions (neutral, happy). The neutral faces were the same as those used in Experiments 2a, 2b, and 3a, but the happy faces differed from those previously used in that no teeth were visible.

(500 trials per task \times 2 tasks \times 40 participants) in Experiments 3a and 3b, respectively. Because individual and group results were very similar for Experiment 3a (with an average correlation of $r = .93$ including both conditions), and because Experiment 3b was based on a large number of observers who each completed a relatively small number of trials, we report group results only.

Mean accuracy in Experiment 3a was similar in the small ($M = 74.48\%$, $SD = 9.79$) and big conditions ($M = 80.36\%$, $SD = 1.84$), $t(2) = 1.227$, $p > .05$. In the small condition, a range of SFs of octave width 1.8 peaking at 5.00 cpf ($Z_{\max} = 12.46$) exceeded the significance threshold ($p < .05$, $S_r = 128$, $FWHM = 3.53$, $Z_{\text{crit}} = 3.25$). The significance threshold (Z_{crit}) is slightly lower than in Experiment 2a because of the reduced stimulus resolution (128×128 instead of 256×256 pixels). In the big condition, an SF range of octave width 2.81 peaking at 8.57 cpf ($Z_{\max} = 9.19$) attained significance. This is a replication of the results of Experiments 2a–c obtained with a similar face width. The difference between the group classification vectors of the two conditions reached significance between 3.57 cpf and 5.00 cpf as well as between 9.29 cpf and 17.86 cpf, with a maximum at 4.29 cpf ($Z_{\max} = 5.27$). Thus, as expected, we observed a clear shift toward lower SFs with smaller stimuli (Figure 10). This shows that the SF Bubbles approach is capable of revealing changes in SF tuning based on the same task and the same stimuli as in Experiment 2a.

In Experiment 3b, mean accuracy across participants was significantly higher for happy/neutral ($M = 81.20\%$, $SD = 5.20$) than for gender discriminations ($M = 65.31\%$, $SD = 5.03$), $t(39) = 16.899$, $p < .001$. The group classification vector results for happy/neutral and gender discriminations are illustrated in Figure 11. We

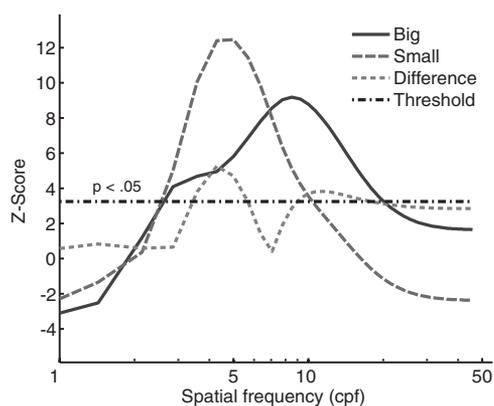


Figure 10. Accuracy group classification vectors for upright face identification obtained in the *big* and *small* conditions of Experiment 3a. The two faces at the bottom show the spatial frequencies (cpf = cycles per face) that reached statistical significance ($Z_{\text{crit}} = 3.25$, $p < .05$) in both size conditions. The difference between classification vectors exceeded the significance threshold.

analyzed the results for the different stimulus types separately, that is, happy male, happy female, neutral male, and neutral female classification vectors for both the happy/neutral and gender discrimination tasks. These eight classification vectors allowed us to compare the use of SF information for the same group of stimuli (e.g., happy females) in the two tasks (gender and happy/neutral) and thus to isolate the top-down effect of task demands.

The group classification vectors for the happy/neutral and gender discriminations revealed different SF patterns for each of the four stimulus types. With happy male faces, the significant portion of the classification vector for the happy/neutral task was shifted into a lower SF range (with peaks at 2.14 cpf, $Z_{\max} = 6.45$, and 5 cpf, $Z_{\max} = 6.37$) than the peak for the gender task (7.86 cpf, $Z_{\max} = 3.94$, $p < .05$, $S_r = 256$, $FWHM = 3.53$, $Z_{\text{crit}} = 3.45$; for details, see Chauvin et al., 2005). The correlation between the classification vectors was $r = -.52$. With neutral male faces, the classification vector for the happy/neutral task peaked at 8.57 cpf ($Z_{\max} = 7.75$), and for the gender task at 10 cpf ($Z_{\max} = 4.10$); the correlation was $r = .44$. With happy female faces, the classification vectors for the two tasks both peaked at 2.86 cpf ($Z_{\max} = 9.27$, happy/neutral; $Z_{\max} = 11.75$, gender) and had a correlation of $r = .90$. With neutral female faces, two distinct peaks were revealed at 7.86 cpf ($Z_{\max} = 9.33$) for happy/neutral and at 2.14 cpf ($Z_{\max} = 7.84$) for gender discriminations, with a correlation between classification vectors of $r = .39$. With all four stimulus types, the difference between the classification vectors for happy/neutral and gender reached significance. Furthermore, the results revealed that for both male and female faces, the happy/neutral classification vectors were only weakly correlated ($r = -.03$ and $r = .22$, respectively), and that for both happy and neutral faces, the gender classification vectors had a relatively weak correlation as well ($r = -.46$ and $r = -.48$). These results are in accordance with the view that the information required for different tasks can reside at different SFs of the same stimulus, and that our visual system is flexibly tuned to extract this information. Overall, the results of Experiment 3b show that SF Bubbles is a technique sensitive to differences in SF tuning that are only due to changes in task demand.

Together the findings of Experiments 3a and b demonstrate that the SF Bubbles approach is capable of revealing subtle differences in SF tuning for complex stimuli induced by a bottom-up factor (Experiment 3a) and by a top-down factor (Experiment 3b). These results also suggest that the null results of Experiments 2a–c are real; thus, the FIE cannot be attributed to qualitative processing differences at the SF level.

General Discussion

The goals of the present study were to introduce a new SF probing technique and to uncover the SFs that mediate upright and inverted face identification. It comprised six experiments: The first experiment was designed to assess the validity of the SF Bubbles technique. Using a plaid detection task, we verified that the SF Bubbles method could reveal precisely the two SFs of the plaid. Compared with traditional SF sampling techniques—such as high-pass, low-pass, and band-pass filtering (e.g., Boutet et al., 2003; Goffaux & Rossion, 2006), or critical band masking (e.g., Gaspar, Sekuler, & Bennett, 2008)—SF Bubbles minimizes the risk that observers adapt to a certain SF range during the task by randomly

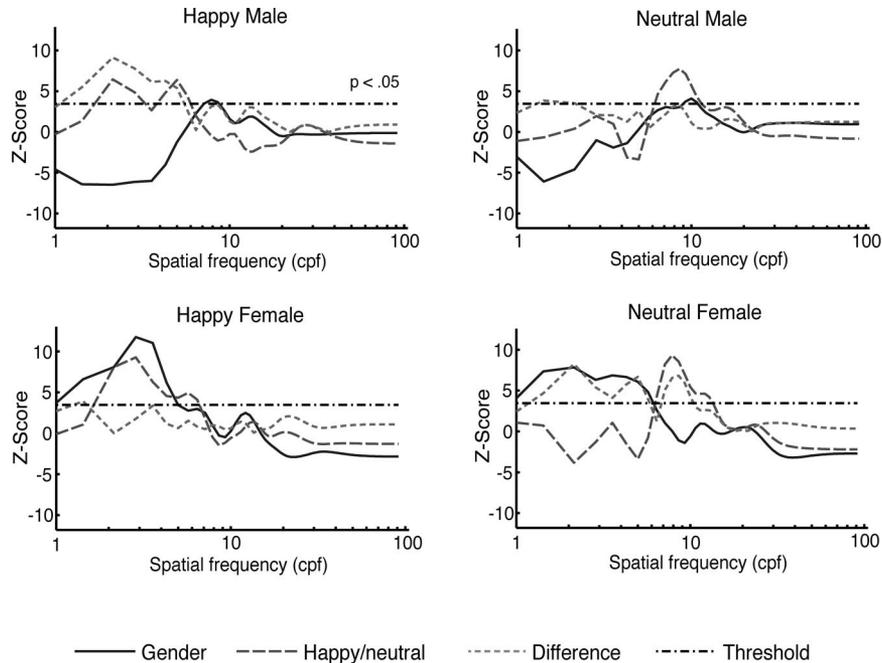


Figure 11. Accuracy group classification vectors obtained for two different tasks (happy/neutral and gender discriminations) performed on the same face set (Experiment 3b). For all stimulus types, the difference between the classification vectors for the two tasks exceeded the significance threshold ($Z_{crit} = 3.45$, $p < .05$).

sampling the SF information on a trial-by-trial basis. Another advantage of SF Bubbles in comparison with critical band masking is that SF Bubbles does not assume that SFs are integrated linearly. We could, in principle, evaluate the joint utilization of two or more SF bands provided that we perform computationally taxing second-order analyses, which go beyond the scope of the current article (e.g., Schyns et al., 2002).

The next three experiments—the main experiments of the article—examined which SFs are diagnostic for the accurate (Experiments 2a and c) and fast (Experiment 2b) identification of upright and inverted faces. Although accuracy was on average 24% higher and RTs were 455 ms shorter with upright faces, thus showing a clear FIE, SF tunings were remarkably similar in both orientation conditions. A single SF band of 1.9 octaves that peaked at 9.8 cpf width was used by observers. This result was obtained independently of whether we used the accuracy (Experiments 2a and c) or RT (Experiment 2b) classification vectors. Moreover, this finding appears to be robust to changes in procedures (many trials and few subjects in Experiments 2a and 2b; few trials and relatively many subjects in Experiment 2c) and to changes in face sets. In Experiment 2c, we employed a subset of the base faces used by Goffaux and Rossion (2006) and replicated our findings of Experiment 2a.

To rule out that this null result was due to an insensitivity of the SF Bubbles technique, Experiment 3a re-examined SF tuning in the identification task of Experiment 2a as a function of stimulus size, which is known to influence SF tuning in a bottom-up fashion (Chung et al., 2002; Loftus & Harley, 2005; Majaj et al., 2002; Näsänen, 1999). In Experiment 3b, task demands were modified (gender or happy/neutral discrimination) to modulate SF tuning in a top-down fashion (e.g., Schyns & Oliva, 1999). In both cases, we revealed subtle differences in SF use, confirming that the SF

Bubbles technique is sensitive to bottom-up- and top-down-induced SF tuning changes. This result suggests that the null results of Experiments 2a–c are real. In addition, the big condition of Experiment 3a, which employed size parameters comparable to those of Experiments 2a–c, closely replicated our previous results.

The SF range revealed in the present study for faces subtending a horizontal visual angle of approximately 6° is consistent with the intermediate-SF band (centered between 7 cpf and 16 cpf) identified as optimal in previous experiments for upright faces of visual angles between 2.3° and 9.5° (e.g., Costen et al., 1994, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold et al., 1999; Näsänen, 1999). Our results are also consistent with those of Boutet et al. (2003), who argued that band-pass filtering faces in the low-SF (1.25–5 cpf), medium-SF (5–20 cpf), or high-SF (20–80 cpf) range had little impact on the FIE. Furthermore, our results are in agreement with a recent study on SF use in upright and inverted face identification by Gaspar, Sekuler, and Bennett (2008). They used critical band noise masking to examine SF tuning in two 10-choice identification tasks and reported that SF tuning for upright and inverted faces was mediated by the same SF band (approximately 1.5 octaves wide and centered at about 7 cpf for face stimuli with a width of 2.3° of visual angle). Our study replicates their findings based on accuracy and complements them by including RT analyses, by using a different SF probing technique, and by using different indexes of the FIE. Specifically, they equated accuracy between orientation conditions by manipulating the quantity of signal, and signal threshold was their index of the FIE. In contrast, we used the same quantity of signal and noise for upright and inverted faces, and response accuracy and RT were our indexes of the FIE.

Our face identification results, however, appear to contradict the findings obtained by Collishaw and Hole (2000) and Goffaux and Rossion (2006), which suggest that inversion might be particularly detrimental to faces containing only low SFs.¹ For example, Goffaux and Rossion reported that the composite face effect (Young et al., 1987) and the whole–part advantage (Tanaka & Farah, 1993)—used as indexes of holistic processing—were more pronounced for low-pass filtered (<8 cpf) than for high-pass filtered (>32 cpf) upright faces. In an additional experiment, they found that the composite effect for upright faces in the low-SF condition was also larger than in an intermediate-SF (8–32 cpf) condition. In contrast, no disproportionate composite effect for low SFs was observed when faces were upside-down. The authors concluded that holistic processing is largely supported by low SFs (but see Cheung, Richler, Palmeri, & Gauthier, 2008, for a re-examination with an extended paradigm). An SF Bubbles experiment is unbiased in comparison with low-pass, high-pass, or band-pass filtering in the sense that on some trials it is equivalent to low-pass filtering, on others to high-pass filtering, and on others to band-pass filtering. On the majority of trials, it is equivalent to band-pass filtering multiple bands simultaneously. At the limit, it contains all possible filtering experiments. If there was a difference in the low-SF range between the identification of upright and inverted faces, this difference should have affected participants' behavior in our experiments at least on the trials in which only low SFs were shown, and we should have seen traces of this in the classification vectors. However, we did not observe such an effect.

How much of the variance between these findings can be explained by different modes of stimulus presentation remains to be investigated. Collishaw and Hole (2002) presented stimuli with both inner and outer facial features (see also Goffaux & Rossion, 2006, Experiment 1). Goffaux and Rossion (2006, Experiments 2–4) showed inner facial features through different elliptical apertures, whereas Gaspar, Sekuler, and Bennett (2008) showed inner facial features through a constant elliptical aperture. We employed the same approach as Gaspar, Sekuler, and Bennett and obtained results consistent with theirs but inconsistent with the studies using a different mode of stimulus presentation.

Another potential explanation for the different findings is that the studies tap into different processes. It is conceivable that holistic processing as indexed by the composite face effect does not correlate with accurate face identification. In fact, Konar, Bennett, and Sekuler (2007, 2008) assessed this possibility by using a composite face task and different identification tasks in a within-subject design. In the composite task, participants were asked to make same–different judgments about the top halves of faces while they were either aligned or misaligned with the bottom halves. In the three identification tasks, participants were asked to either determine whether a target face was present in a lineup, or perform a 10-alternative forced-choice identification with unlimited viewing time, or perform a 4-alternative forced-choice identification with a viewing time of 200 ms (i.e., the viewing duration in the composite task). No correlation was found between the strength of the composite face effect and accuracy in any of these face identification tasks. If there is a qualitative difference between the processing of upright and inverted faces—possibly in SF use—that is neither correlated with accuracy nor with RT in face identification, it will not be revealed in our classification vectors. Thus, even though we did not find differences in SF use for the

identification of upright and inverted faces, our results are not necessarily inconsistent with a qualitative view of the FIE.

It also remains possible that the same SFs are used, but in a different fashion. For example, in the upright condition, observers might use information at 9.8 cpf to encode the distance between eye and eyebrow; in the inverted condition, they might use information at the same SF to encode local features of the eye. Accordingly, Boutet et al. (2003) showed that intermediate SFs are optimal for both configural and featural modifications (but see Goffaux, Hault, Michel, Vuong, & Rossion, 2005, for a dissociation between low- and high-SF information and the extraction of configural and featural cues). Although the present study provides a precise estimate of the SF tuning curves and demonstrates that the performance drop with inversion is not due to a shift to less informative SFs, it remains an avenue for future research to shed light on how exactly the revealed SF information is used at other processing stages.

The most parsimonious explanation for the present findings is provided by the quantitative account of the FIE (Sekuler et al., 2004). According to this view, the FIE can be explained in terms of a decrease in the sensitivity of the same process. Sensitivity can be broken down into *calculation efficiency*, its deterministic component, and *internal noise*, its stochastic component. Gaspar, Bennett, and Sekuler (2008) recently found evidence that face inversion leads to a reduction in calculation efficiency but does not alter internal noise. We interact much more frequently with upright than with inverted faces, and it is thus plausible that a neural mechanism similar to that thought to mediate *perceptual learning* is responsible for the FIE. For instance, Kobatake, Wang, and Tanaka (1998) showed that monkeys trained on a set of novel stimuli have more inferotemporal neurons responsive to features of these stimuli than untrained monkeys. The more such selective neurons, the greater is the sensitivity. Perhaps more fusiform gyrus neurons are responsive to upright than to inverted facial features (see also Perrett, Oram, & Ashbridge, 1998, for a review) within the critical SF band characterized in this article.

Conclusion

The present study introduced a novel SF sampling technique that was applied to investigate SF tuning in upright and inverted face identification. The results show that the same SFs (dependent on stimulus size) were used for the accurate and fast identification of upright and inverted inner facial features, and performance was higher in the upright condition. The findings place an additional constraint on theories of qualitative processing differences and are consistent with predictions of the quantitative account of the FIE.

¹ Nakayama (2003) reported a broader masking function for inverted than for upright faces, which is also in disagreement with our SF results. To our knowledge, this study has been published as an abstract only, and we do not have sufficient details to discuss it any further.

References

- Ahumada, A., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751–1756.

- Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, *25*, 281–316.
- Barton, J. J., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology*, *92*, 527–549.
- Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception*, *35*, 1089–1105.
- Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics*, *65*, 1078–1093.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*, 659–667.
- 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.
- Chung, S. T. L., Legge, G. E., & Tjan, B. S. (2002). Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Research*, *42*, 2137–2152.
- 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.
- Collishaw, S. M., & Hole, J. C. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception*, *29*, 893–909.
- Costen, N. P., Parker, D. M., & Craw, I. (1994). Spatial content and spatial quantisation effects in face recognition. *Perception*, *23*, 129–146.
- Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception and Psychophysics*, *58*, 602–612.
- Davies, G., Ellis, H., & Shepherd, J. (1977). Cue saliency in faces as assessed by the “photofit” technique. *Perception*, *6*, 263–269.
- De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision*. New York: Oxford University Press.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.
- Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision*, *2*, i–i. doi:10.1167/2.1.i
- Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 628–634.
- Fiset, D., Blais, C., Gosselin, F., & Schyns, P. (2006). Effective frequency tuning of three face categorization tasks [Abstract]. *Journal of Vision*, *6*, 273, 273a. doi:10.1167/6.6.273
- Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, *29*, 159–170.
- 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.
- Gaspar, C. M., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research*, *48*, 2817–2826.
- 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. M., 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.
- Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, *28*, 167–207.
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, *41*, 2261–2271.
- Gosselin, F., & Schyns, P. G. (2005). Bubbles: A user’s guide. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 91–106). Hillsdale, NJ: Erlbaum.
- Haig, N. D. (1985). How faces differ—A new comparative technique. *Perception*, *14*, 601–615.
- Haig, N. D. (1986). Exploring recognition with interchanged facial features. *Perception*, *15*, 235–247.
- Hochberg, J., & Galper, R. R. (1967). Recognition of faces: An explanatory study. *Psychological Science*, *9*, 619–620.
- Kobatake, E., Wang, G., & Tanaka, K. (1998). Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *Journal of Neurophysiology*, *80*, 324–330.
- Konar, Y., Bennett, P. J., & Sekuler, A. B. (2007). The composite face effect is not correlated with face identification accuracy [Abstract]. *Journal of Vision*, *7*, 501, 501a. doi:10.1167/7.9.501
- Konar, Y., Bennett, P. J., & Sekuler, A. B. (2008). The composite face effect is still not correlated with face identification accuracy [Abstract]. *Journal of Vision*, *8*, 891, 891a. doi:10.1167/8.6.891
- Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *51(A)*, 449–473.
- Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *53(A)*, 513–536.
- Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception*, *30*, 73–83.
- Leder, H., & Carbon, C.-C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology*, *97*, 19–29.
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature*, *410*, 890.
- Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin and Review*, *12*, 43–65.
- Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, *42*, 1165–1184.
- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.
- McCotter, M., Gosselin, F., Sowden, P., & Schyns, P. G. (2005). The use of visual information in natural scenes. *Visual Cognition*, *12*, 938–953.
- 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.
- Nakayama, K. (2003). Face specific processing: Role of local features in an affine metric [Abstract]. *Journal of Vision*, *3*, 91, 91a. doi:10.1167/3.9.91
- Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, *39*, 3824–3833.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumula-

- tion in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition*, 67, 111–145.
- Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception*, 17, 43–63.
- Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, 47, 25–57.
- Riesenhuber, M., Jarudi, I., Gilad, S., & Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B: Biological Sciences*, 271, S448–S450.
- Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces? *Cognition*, 88, 79–107.
- Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, 103, 34–79.
- Rock, I. (1974). The perception of disoriented figures. *Scientific American*, 230, 78–85.
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, 128, 274–289.
- Rossion, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal of Vision*, 8, 1–13.
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, 1, 62–74.
- 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., 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.
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, 14, 391–396.
- Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, 75, 221–242.
- Shepherd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 105–131). London: Academic Press.
- Sowden, P. T., & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Sciences*, 10, 538–545.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 46(A), 225–245.
- Tanaka, J. W., & Sengco, J. (1997). Features and their configuration in face recognition. *Memory and Cognition*, 25, 583–592.
- Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception*, 9, 483–484.
- Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79, 471–491.
- Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, 6, 313–326.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics*, 33, 113–120.
- Williams, C. C., & Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory and Cognition*, 35, 1977–1985.
- Yarmey, A. D. (1971). Recognition memory for familiar “public” faces: Effects of orientation and delay. *Psychonomic Science*, 24, 286–288.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141–145.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16, 747–759.
- Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, 44, 889–898.

Received April 15, 2008

Revision received March 5, 2009

Accepted March 7, 2009 ■