

Behavioral Evidence for a Predominant and Nonlateralized Coarse-to-Fine Encoding for Face Categorization

Rui de Moraes Jr.

University of São Paulo and University of Montreal

Louise Kauffmann

Grenoble Alpes University and Max Planck
Institute for Human Cognitive and Brain Sciences,
Leipzig, Germany

Sérgio S. Fukusima

University of São Paulo

Jocelyn Faubert

University of Montreal

Influential models on visual perception assume that there is a precedence of low over high spatial frequencies (SFs) in the processing time course of the visual input, that is, a coarse-to-fine (CtF) encoding. Additionally, hemispheric asymmetries for strategies of SF processing have been shown. A CtF processing would be favored in the right hemisphere, whereas the reverse fine-to-coarse (FtC) processing would be favored in the left hemisphere. In the current article, we aimed to behaviorally investigate which temporal strategy, that is, CtF or FtC, each brain hemisphere performs to integrate SF information of human faces. To address this issue, we conducted a male–female categorization task using the divided visual field paradigm; CtF and FtC brief dynamic sequences of faces were presented in the left, right, and central visual fields. Results of the correct response time and the inverse efficiency score showed an overall advantage of CtF processing for face categorization, irrespective of the visual field of presentation. Error rate data also highlights the role of the right hemisphere in CtF processing. Here, we provide evidence at the behavioral level for a general and nonlateralized precedence of the default CtF strategy carried out by the visual system to encode faces, a complex stimulus with ecological value.

Keywords: face perception, spatial frequency, hemispheric specialization, coarse-to-fine, fine-to-coarse

The human face is a special class of visual stimulus because of its biological and social relevance. Evolutionary and environmental pressures pushed the visual system to develop efficient strategies to encode facial information. Thus, there is a whole field devoted to under-

standing how visual information is processed and integrated during face perception. From a neurobiological point of view, neurophysiological data indicate that cells of the primary visual cortex decompose the visual input mainly in terms of spatial frequencies (SFs; i.e., periodic

This article was published Online First October 20, 2016. Rui de Moraes, Jr., Department of Psychology, FFCLRP, University of São Paulo and School of Optometry, University of Montreal; Louise Kauffmann, Grenoble Alpes University and Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; Sérgio S. Fukusima, Department of Psychology, FFCLRP, University of São Paulo; Jocelyn Faubert, School of Optometry, University of Montreal.

Rui de Moraes, Jr., was supported by the State of São Paulo Research Foundation (FAPESP) under Grant 2012/00945-2, and a Natural Sciences and Engineering Research Council (NSERC)

discovery fund to Jocelyn Faubert. Louise Kauffmann was supported by Région Rhône-Alpes (Cible Grants).

Correspondence concerning this article should be addressed to Rui de Moraes, Jr., Departamento de Psicologia – FFCLRP, Universidade de São Paulo, Av. Bandeirantes, 3900, CEP 14040-901, Ribeirão Preto, SP, Brazil, or to Louise Kauffmann, Laboratoire de Psychologie et Neurocognition, Université Pierre Mendès France, BP 47, 38040 Grenoble Cedex09, France. E-mail: ruidemoraesjr@gmail.com or louise.kauffmann@gmail.com

variations of luminance through space) and orientations. More and more complex computations are subsequently performed by higher level areas along the ventral visual stream, until visual information is integrated to yield high-level face representations in inferior temporal cortices (de Valois, Albrecht, & Thorell, 1982; Hubel & Wiesel, 1968; Poggio, 1972). However, it is not clear exactly how sensorial low-level information such as SF is integrated in high-order cognitive representations during face perception.

Influential and recent models of visual perception assume that the visual system integrates SF information following a predominantly coarse-to-fine (CtF) processing strategy (Bar, 2003, 2007; Bar et al., 2006; Bullier, 2001; Hegdé, 2008; Kauffmann, Ramanoël, & Peyrin, 2014; Marr, 1982; Schyns & Oliva, 1994). According to these models, low SFs, which convey coarse information about a visual stimulus, are rapidly processed via fast magnocellular pathways, enabling a coarse parsing of the visual input. This initial low-pass analysis is subsequently refined by the extraction of detailed information contained in high SFs, conveyed more slowly through parvocellular pathways. Evidence of such CtF processing has been found in studies using a wide variety of visual stimuli: sinusoidal gratings (Breitmeyer, 1975), hierarchical forms (Navon, 1977), hybrid images (Schyns & Oliva, 1994), natural scenes (Kauffmann, Chauvin, Guyader, & Peyrin, 2015; Musel, Chauvin, Guyader, Chokron, & Peyrin, 2012), and human faces (Z. Gao & Bentin, 2011; Goffaux et al., 2011; Halit, de Haan, Schyns, & Johnson, 2006; Vlaming, Goffaux, & Kemner, 2009). Therefore, the literature supports the CtF strategy as a general and default integration mode performed by the visual system to process SF information.

However, many studies showed that the strategy of SF processing is flexible and that a CtF processing strategy can be reversed in favor of a fine-to-coarse (FtC) strategy, depending on factors such as task constraints (e.g., Oliva & Schyns, 1997; Schyns & Oliva, 1994; Schyns & Oliva, 1999) and hemispheric specialization. For example, Peyrin et al. (2005) conducted an fMRI study to investigate how the brain hemispheres integrate SF information during scene perception. Pairs of scene images were presented in succession for 100 ms and separated by a time gap of 400 ms. The pairs of images

were presented in either a CtF (a low-pass scene followed by a high-pass scene) or FtC sequence (a high-pass scene followed by a low-pass scene). For fMRI data analysis, the authors used a direct interhemispheric comparison method to address hemispheric asymmetries during CtF and FtC scene processing. This method enables the assessment of whether activity in regions of one hemisphere differs from activity in homologous regions of the opposite hemisphere according to the experimental conditions. Results showed that CtF sequences elicited greater activation within the right than left occipito-temporal cortex, whereas FtC sequences elicited greater activation in the left than right occipito-temporal cortex. These results thus suggested a right-hemispheric predominance for CtF processing and a left-hemispheric predominance for FtC processing.

These results are supported by other studies suggesting a functional brain asymmetry to process SF information. Sergent (1982) first proposed that the right hemisphere is more efficient in processing low SFs, whereas the left hemisphere is more efficient in processing high SFs. A large empirical framework further supported this hypothesis (e.g., dos Santos, Andrade, & Fernandez Calvo, 2013; Musel et al., 2013; Reinvang, Magnussen, & Greenlee, 2002). Interestingly, past studies also showed SF asymmetry for face processing (de Moraes, Sousa, & Fukusima, 2014; Perilla-Rodríguez, de Moraes, & Fukusima, 2013; Sergent, 1985). Overall, previous studies on SF processing and hemispheric functional asymmetries led us to wonder about the temporal integration of SF to encode human faces in the brain hemispheres.

Despite the general assumption that SF processing is time-dependent, most of the past studies on SF processing during face perception used long stimuli exposure duration or did not directly address the CtF hypothesis (Eger, Schyns, & Kleinschmidt, 2004; Gauthier, Curby, Skudlarski, & Epstein, 2005; Iidaka, Yamashita, Kashikura, & Yonekura, 2004; Vuilleumier, Armony, Driver, & Dolan, 2003). Some insights about this issue were provided in a recent fMRI study using a slow event-related design conducted by Goffaux et al. (2011). In this study, participants viewed intact- or scrambled-phase versions of filtered faces containing low, middle, or high SFs presented for 75, 150, or 300 ms and subsequently masked. They per-

formed an intact-scrambled manual categorization task. Face-preferring areas (bilateral fusiform face area [FFA], superior temporal sulci, anterior inferotemporal cortex, and right occipital face area) were identified with an independent functional localizer. Data showed that the bilateral FFA, superior temporal sulci, and anterior inferotemporal cortex robustly responded more strongly to low SFs at 75 ms. This response to low SFs decayed with face exposure duration, mostly at the 150-ms condition. Conversely, in the 150-ms or 300-ms condition (depending on the face-sensitive site), response to high SFs increased. These results therefore suggested that a CtF integration of SF information was favored within face-selective cortex. Importantly for our purpose, the results did not reveal any hemispheric asymmetry in the strategy of SF processing for face encoding within these regions. However, these authors did not directly compare activity in homologous face-selective regions of each hemisphere according to the experimental conditions, as performed in the study of [Peyrin et al. \(2005\)](#). It is therefore not guaranteed that both hemispheres process SF information in a predominantly CtF fashion during face processing.

The present study aimed to examine, at the behavioral level, the strategy of SF processing favored during face categorization, and explicitly considered potential hemispheric asymmetries in this process. A behavioral experiment is interesting because the relatively low temporal resolution (i.e., sampling rate) of fMRI may impair investigations on the processing time course of rapid sensorial and cognitive operations ([Amaro & Barker, 2006](#)). In addition, neurophysiological data do not always represent the behavioral output.

In order to investigate strategies of SF processing during face encoding, we used sequences of filtered faces adapted from previous studies on scene perception ([Kauffmann et al., 2015](#); [Musel et al., 2012](#)). These sequences mimic the default CtF or the reverse FtC sequence of SF processing, in order to impose either of these two processing strategies while participants performed a categorization task. In order to address hemispheric asymmetries, these stimuli were presented in a divided visual field. This classic experimental paradigm enables behavioral investigations on functional brain asymmetry. It consists of brief lateralized

presentations of the stimulus. The anatomical structure of the visual system validates the approach, as the right hemisphere initially processes a stimulus projected in the left visual field (LVF), and the left hemisphere initially processes a stimulus projected in the right visual field (RVF; see [Bourne, 2006](#)).

We used the divided visual field method to present SF filtered faces in CtF and FtC brief sequences in the LVF, RVF, and central visual field (CVF). The participants performed a male–female categorization task while response latency and accuracy were recorded. Our task aimed to behaviorally tackle high-level vision. The male–female categorization is a real-world-based task with ecological value, and it is more complex than the “intact vs. scrambled” task in [Goffaux et al. \(2011\)](#); although it was well-suited for the fMRI design). The displayed sequences, besides simulating the visual system operation, samples more SFs than most studies, which generally use a single filter for each extreme in the SF domain.

Based on previous data suggesting a predominant CtF processing strategy for face perception ([Goffaux et al., 2011](#)), we expected that stimuli depicting a CtF sequence would be categorized faster overall than those depicting a FtC sequence in a face categorization task. Furthermore, if the strategy of SF processing for face encoding varies according to the brain hemisphere predominantly involved to perform the task, we would expect better performances to categorize CtF and FtC sequences presented in the LVF and in the RVF, respectively.

Method

Participants

Thirty-two right-handed students with normal or corrected-to-normal vision were recruited at the University of São Paulo. Two students failed the acuity test as assessed by a Snellen chart (visual acuity below 6/7.5), and another was considered ambidextrous when evaluated by the Edinburgh Inventory ([Oldfield, 1971](#)). Thus, 29 subjects (16 females) took part in the study (mean age = 22.9 years, $SD = 4.6$; mean score in Edinburgh Inventory = 84.7, $SD = 18.8$). None of them had neurological or ocular disorders. All participants read and signed the Statement of Consent approved by the local Research Ethics Committee.

Stimuli

Forty images of faces (half females) posing in frontal view and with neutral expressions were extracted from the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist, Flykt, & Öhman, 1998). Facial cues that could enhance the contribution of specific SFs to the task (e.g., wrinkles, blemishes, pimples, scars) were removed using Photoshop 7.0 (Adobe Systems Inc., San Jose, CA, USA). The faces were gray-scale transformed and inserted into a 256×256 pixel-size quadrant. Viewed at 85 cm, the images subtended a visual angle of 5.8×5.8 degrees and the faces themselves encompassed about 4.0 degrees of visual angle. An egg-shaped mask in uniform medium gray superimposed external features (e.g., hair, ears, neck).

The filtering process was performed using MATLAB 7.9.0 (Mathworks Inc., Sherborn, MA, USA) as implemented by Kauffmann et al. (2015). Each image was filtered by six Gaussian bandpass filters with central frequencies set at 4, 5.6, 8, 12, 17, and 24 cycles per image (cpi), with a standard deviation of 4.3 cpi, which corresponds to 0.7, 1, 1.4, 2, 2.9, and 4.1 cycles per degree (cpd), with a standard deviation of 0.7 cpd. We removed SF information below 1.2 cpi (0.2 cpd) and normalized the luminance (i.e., mean luminance of 128 on a 256 gray-level scale) among the filtered faces. The central frequencies used to create the stimuli composing the image sequences were based on a previous study by Kauffmann et al. (2015). These central frequencies encompass SF ranges critical to different tasks and models of processing identified in previous studies on face perception literature (e.g., Boutet, Collin, & Faubert, 2003; Parker & Costen, 1999; Schyns & Oliva, 1999; Watier, Collin, & Boutet, 2010). The central frequencies of the bandpass filters were not linearly spaced but followed a logarithmic scale.¹ This was done in order to obtain a better sampling of the amplitude spectrum of natural stimuli such as faces, in which amplitude decreases as SF increases (“ $1/f$ ” shape; Field, 1987), and a better sampling of low SFs with more filters centered on low SFs (see Kauffmann et al. [2015] for more details on the filtering procedure, and for a similar approach, see Willenbockel et al. [2010]).

We used the filtered versions of the stimuli to create brief dynamic sequences. The six filtered versions of each face were assembled to create a sequence. They were presented in succession

going from lower to higher SFs (CtF sequences), or vice versa (FtC sequences). Thus, only the order of presentation differentiated CtF from FtC sequences. The sequences lasted 141 ms, and each one of the six filtered faces was displayed for an average time of 23.5 ms (refresh rate set at 85 Hz) on a 19-in. CRT monitor. Figure 1 shows examples of the stimuli used.

Procedure and Experimental Design

The experimental procedure was performed in an individual and single session in a dark and adapted room. Instructions were given by the researcher and the computer screen, and emphasized the importance of fixating at the central fixation point during the stimulus presentation.

Each trial began with the presentation of a central fixation point for 500 ms in order to drive the participant’s gaze to the screen center. It was immediately followed by a dynamic sequence lasting 141 ms. The sequences could be presented either in the LVF, RVF, or CVF, in a CtF or FtC succession. When the stimulus was presented lateralized, the face’s inner edge was 2.5° of visual angle distant from the fixation point, which was still displayed on the screen. In addition, the opposite hemifield was filled by a mask at the same size and eccentricity of the stimulus and presented for the same time period. This procedure improves fixation control over trials by avoiding that attention driven to a unilaterally presented stimulus initiates a saccade toward it² (Carpenter, 1988). Next, the same mask covered the stimuli area for 35 ms to

¹ The logarithmic scale used to define central frequencies of the filters was given by $f_k = f_o/X^k$, where f_o is the highest central frequency, X is the geometric separation between each filter central frequency, and k is the filter number from high to low frequency.

² Eye movements were not recorded in this experiment. However, it is possible to safely rely on the experiment output when employing the appropriate controls even when an eye-tracking device is not available (see Bourne, 2006). Our experiment controlled saccadic-related variables such as eccentricity, degrees of visual angle of stimulus width, randomized hemifield presentation, fixation cue, and exposure time. In particular, the exposure time plays a key role in saccade occurrence. Mean saccadic latencies range from 150 ms to 200 ms, and only 2% of the latencies occur before 150 ms (Carpenter, 1988). We conducted a follow-up study monitoring the gaze location that support our view; SF filtered faces in a divided visual field were presented using very similar parameters to the present study, and around 5% of the trials were invalidated due inaccurate gaze, which is a quite acceptable rate.

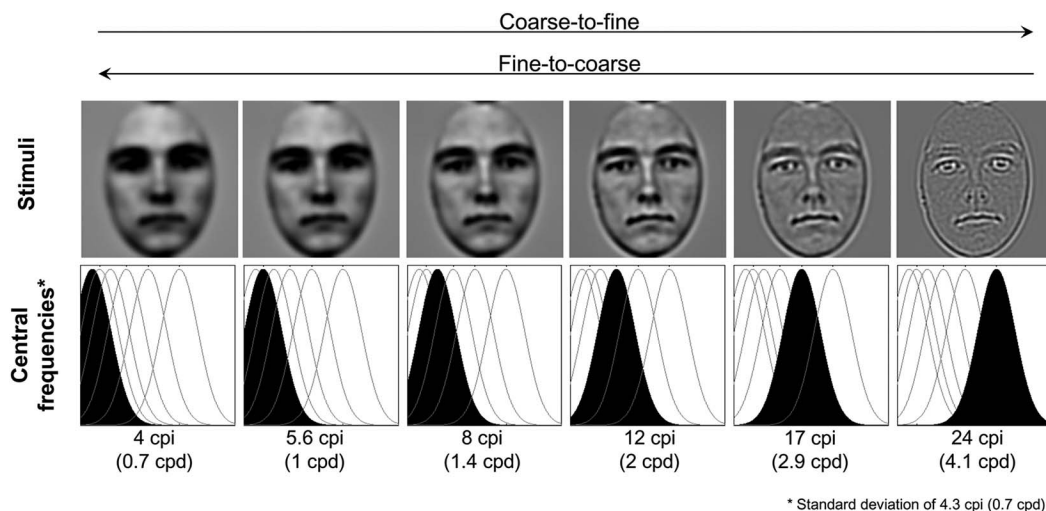


Figure 1. Example of the six spatial-frequency filtered versions that originated from a full-bandwidth image along with the information of their central frequency measured in cycles/image width (cpi) and cycles/degree of visual angle (cpd). One stimulus consisted of a brief succession of the filtered faces going from lower (left side) to higher (right side) spatial frequencies in the coarse-to-fine condition, or in the opposite direction going from higher to lower spatial frequencies in the fine-to-coarse condition (KDEF stimulus ID: AM08NES, with permission from Karolinska Institute).

prevent retinal persistence. The masks were built in order to have a $1/f$ amplitude spectrum, where f is the SF, and a random phase spectrum to match the amplitude spectrum of natural stimuli such as faces, in which amplitude decreases with SF following a $1/f$ function³ (Field, 1987; Tolhurst, Tadmor, & Chao, 1992). Thus, we ensured that the amplitude spectrum of the mask was similar to the mean amplitude spectrum of the face stimuli, and masked low SFs as efficiently as high SFs. At the mask offset, participants had to categorize the face as male or female as fast and as accurately as possible. Figure 2 illustrates one trial. The answer was given by pressing a blue or red button with the index or middle finger, depending on the response category (button positions were counter-balanced across participants). An RB-730 response pad (Cedrus Corp., San Pedro, CA, USA) registered the response and its latency during a maximum time gap of 1,500 ms. The following trial started just after the response.

All faces from the original stimuli set were presented twice in each condition. Thus, the 3 (Visual Field) \times 2 (Sequences) design had 80 trials per experimental condition, totaling 480 trials, randomly presented within and among

conditions, plus 36 training trials at the beginning of the experiment. Stimuli used for training were not part of the experiment. The entire experiment lasted approximately 15 min. E-Prime (Psychology Software Tools Inc., Pittsburg, PA, USA) displayed the stimuli and collected the data.

Results

We first analyzed the error rate (ER) and the response time (RT) for categorizing the faces. For the RT analysis, only correct responses and latencies inside the boundaries of two standard deviation units of the average correct RT in each condition for each participant were consid-

³ In fact, the amplitude spectra of natural images, such as faces or landscapes, decays as a function of SF falling in a form $1/f^\alpha$, with α ranging from 0.7 to 2 (Field, 1987; Tolhurst et al., 1992). More importantly, the amplitude spectra of the KDEF database follows a form $1/f^{1.61}$ (Wu, Xu, Dayan, & Qian, 2009). It is thus possible that some face stimuli actually contained more low SFs than the mask, which could have resulted in a less efficient masking of low SFs in these stimuli for the FtC condition. Thus, adopting $\alpha = 1.61$ in future studies when using the KDEF database could improve backward masking, especially when it comes to low SFs.

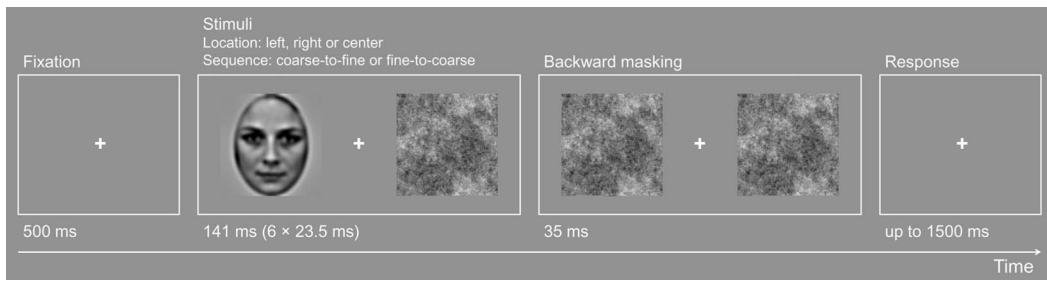


Figure 2. Illustration of one trial. A fixation screen was followed by the stimulus—a coarse-to-fine or fine-to-coarse dynamic sequence that was presented in the left, right, or central visual field and immediately backward-masked. At the mask offset, participants must categorize the face as male or female as fast and as accurately as possible (KDEF stimulus ID: AF01NES, with permission from Karolinska Institute).

ered. This procedure excluded 19.89% trials for judgment errors and omissions and 3.67% for extreme values, totaling 23.56% of the overall trials. A two-way repeated-measures ANOVA ($\alpha = .05$) was performed with Visual Field (LVF, RVF, and CVF) and Sequence (CtF and FtC) as within-participant factors for both ER and RT. We used the Bonferroni adjustment for multiple comparisons. Statistical analyses were conducted with SPSS PASW 18 (IBM – SPSS Inc., Hong Kong, China). **Figures 3a** and **3b** show the mean ER and the mean RT along standard errors of the mean for each experimental condition.

The ANOVA performed on ER revealed a significant main effect of Visual Field, $F(2, 56) = 62.33, p < .001, \eta_p^2 = .69$. Participants made less errors when the faces were presented in the CVF (mean \pm standard error [*SE*] = $12.63\% \pm 1.05\%$) compared with lateralized presentations (LVF = $23.43\% \pm 1.60\%$; RVF = $22.59\% \pm 1.64\%$; $p < .001$ in both comparisons). There was no main effect of Sequence, $F(1, 28) = 1.01, p = .324, \eta_p^2 = .04$. However, there was an interaction between Visual Field and Sequence, $F(2, 56) = 3.49, p = .037, \eta_p^2 = .11$. Further comparisons revealed that when stimuli were presented in the LVF, participants were more accurate in categorizing faces in a CtF than FtC sequence (CtF = $22.03\% \pm 1.77\%$; FtC = $24.83\% \pm 1.64\%$; $p = .027$), whereas no difference was found between CtF and FtC sequences when stimuli were displayed in the RVF (CtF = $22.37\% \pm 1.66\%$; FtC = $22.80\% \pm 1.82\%$; $p = .711$) or in the CVF (CtF = $13.32\% \pm 1.09\%$; FtC = $11.94\% \pm 1.20\%$; $p = .154$).

The ANOVA performed on RT also revealed a significant main effect of Visual Field, $F(2, 56) = 23.39, p < .001, \eta_p^2 = .46$. Participants categorized the faces more rapidly when they were presented in the CVF ($405 \text{ ms} \pm 15 \text{ ms}$) than in the LVF ($439 \text{ ms} \pm 17 \text{ ms}$) and RVF ($439 \text{ ms} \pm 18 \text{ ms}$; $p < .001$ in both comparisons). There was a marginal main effect of Sequence, $F(1, 28) = 3.89, p = .058, \eta_p^2 = .12$, favoring categorization of CtF ($425 \text{ ms} \pm 16 \text{ ms}$) over FtC ($430 \text{ ms} \pm 17 \text{ ms}$) sequences. The interaction between Visual Field and Sequence was not significant, $F(2, 56) = 0.01, p = .991, \eta_p^2 < .01$.

We also analyzed the inverse efficiency score (IES), proposed by [Townsend and Ashby \(1978\)](#). The IES combines the ER and the RT in a single dependent variable. Besides summarizing behavioral findings, the IES circumvent speed–accuracy trade-offs and individual differences in strategy (e.g., one is concerned with just speed or accuracy). The IES is computed individually per condition as follows: $\text{IES} = \text{RT}/(1 - \text{ER})$, expressed in milliseconds. **Figure 3c** shows the mean IES and the standard error of the mean for each experimental condition.

The IES received the same statistical treatment as the ER and RT. The ANOVA performed on IES revealed a significant main effect of Visual Field, $F(2, 56) = 67.91, p < .001, \eta_p^2 = .71$. As expected, the efficiency in categorizing faces was greater in the CVF (CVF = $464 \text{ ms} \pm 16 \text{ ms}$; LVF = $577 \text{ ms} \pm 22 \text{ ms}$; RVF = $571 \text{ ms} \pm 24 \text{ ms}$; both comparisons with $p < .001$). The main effect of Sequence reached significance, $F(1, 28) = 4.23, p = .049, \eta_p^2 = .13$, showing that face categorization was

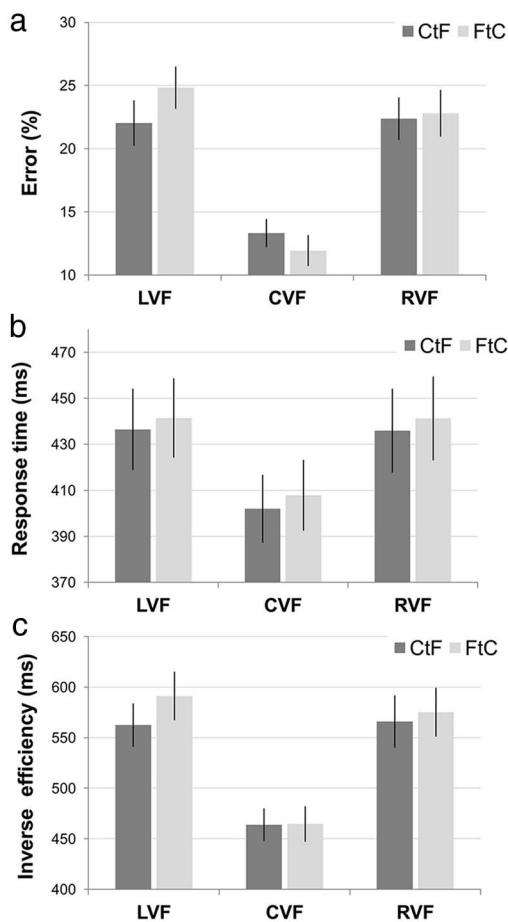


Figure 3. Averages of (a) error rate, (b) correct response time, and (c) inverse efficiency score for dynamic coarse-to-fine (CtF) and fine-to-coarse (FtC) sequences of faces presented in the left (LVF), central (CVF), and right visual field (RVF). Error bars indicate the standard error of the mean.

more efficient in CtF ($531 \text{ ms} \pm 20 \text{ ms}$) than FtC ($544 \text{ ms} \pm 20 \text{ ms}$) sequences. The interaction between Visual Field and Sequence was not significant, $F(2, 56) = 1.98, p = .148, \eta_p^2 = .07$ (Greenhouse-Geisser corrected, $\epsilon = .80$).

In short, the analysis of ER showed an interaction: When presentation occurred in the LVF, categorizations of CtF sequences were more accurate than categorization of FtC sequences, whereas there was no difference between categorization of CtF and FtC sequences when presented in the CVF or RVF. Results regarding the RT in the LVF showed the same pattern, but

with no significant difference, as is shown by the flattened bars in Figure 3b. Both IES and RT analysis evidenced no interaction between the Sequence and Visual Field factors. However, IES and RT showed a significant effect and a strong trend toward significance, respectively, for the main factor of Sequence, with an overall advantage for CtF processing irrespective of the brain hemispheres.

Discussion

The present behavioral study aimed to provide supplementary arguments in favor of a predominantly CtF processing strategy carried out by the visual system to encode faces. We additionally examined whether this predominant CtF processing strategy could be reversed in favor of a FtC processing strategy, depending on the brain hemisphere used to perform the task. For this purpose, we implemented the divided visual field method. Our results, as measured by RT and IES, showed an overall better efficiency in categorizing faces (male vs. female categorization) in a CtF than in a reverse FtC sequence, regardless of the visual field of presentation.

First, this overall CtF advantage suggests that the visual system initially extracts low SFs conveyed by fast magnocellular pathways and builds a coarse face representation. This representation is later on refined by high SFs conveyed by slower parvocellular pathways. These data are consistent with a previous fMRI study conducted by Goffaux et al. (2011), wherein most face-sensitive sites produced a CtF activation in the brain hemispheres. Therefore, the previously mentioned study and our data agree that temporal integration of SF for face encoding is primarily unidirectional and stable across the brain hemispheres. Therefore, there is no evidence of an inversion in the temporal processing of SF in the left hemisphere as was found by Peyrin et al. (2005). In that study, FtC sequences of scenes revealed greater activation within the left occipito-temporal cortex; the right side revealed the default CtF preference.

Previous investigations on face encoding that did not consider functional brain asymmetry support the CtF model. A previous behavioral investigation also found a CtF pattern for face categorization using fixed temporal constraint conditions and a psychophysical response vari-

able rather than dynamic sequences and response latencies. Participants performed a same-different face recognition task. Pairs of faces were presented simultaneously, one above the other, for 250, 500, or 800 ms. Vertical alignment prevented hemispheric differences. One of the faces contained a broad range of SFs, and the other contained high or low SFs. The A_z signal detection index was greater for recognition of low-SF faces when compared with high-SF faces at the higher temporal constraint condition, 250 ms. However, results showed no difference in the 500-ms and 800-ms exposure time conditions (Z. Gao & Bentin, 2011, Experiment 1). Similarly, electrophysiological studies report a CtF order of encoding. An experiment implementing a passive paradigm presented fearful and neutral faces containing high and low SFs for 500 ms, with a random interstimulus interval (1,600 to 1,800 ms). Results of the P1 and N170 ERPs showed that besides playing a core role in brain response to fear, low SFs are processed before high SFs for face encoding (Vlamings et al., 2009, Experiment 1). In another ERP study, famous faces were filtered in four SF bands: high SF, low SF, high-and-low SF by superimposition, and high-and-low SF scrambled noise as control. Stimuli were presented during 120 ms with a random interstimulus interval (1,000 to 2,000 ms) in a background that maintained similar spectral energy among conditions. Participants performed a yes-no face detection task. Results of the N170 for low-SF faces showed larger amplitude and quicker peak latency when compared with high-SF faces (Halit et al., 2006). Altogether, these studies and our results suggest that face categorization follows a CtF integration order.

As a secondary result, the ER data showed that when presentation occurred in the LVF, participants were more accurate in categorizing CtF sequences. It suggests that the right hemisphere predominantly performs this strategy. This CtF advantage might be related to the holistic processing performed by the right hemisphere to encode faces (Jacques & Rossion, 2015; Maurer et al., 2007; Renzi et al., 2013; Rossion et al., 2000). Holistic processing emerges very early during face processing (Ramon & Rossion, 2012; Richler, Mack, Gauthier, & Palmeri, 2009) and relies on low SFs (Collishaw & Hole, 2000; Gof-

faux, 2009; Goffaux, Hault, Michel, Vuong, & Rossion, 2005; Goffaux & Rossion, 2006; however, see Collin, Rainville, Watier, & Boutet, 2014). Therefore, taking together the results of RT, IES, and ER, we can assume that both brain hemispheres preferentially integrate SF information in a CtF order, and that this process may be more pronounced in the right hemisphere.

Outside the scope of face perception, the present study supports influential and general CtF models on visual perception (e.g., Bar, 2003; Bullier, 2001; Marr, 1982). Here, we provide a behavioral contribution to the field by issuing laterality effects. In addition, we used a complex stimulus, the human face. Besides its biological and social relevance, the human face advantageously seems to be more sensitive to SF than other visual stimuli. Selective extraction of SF facial information impairs the execution of specific tasks more markedly than for most visual stimuli. For example, low SFs are essential for configural representation, facial identity relies on intermediate SFs, and perception of local elements is based on high SFs (Collin, Liu, Troje, McMullen, & Chaudhuri, 2004; Goffaux, Gauthier, & Rossion, 2003; Yue, Tjan, & Biederman, 2006).

However, our assumptions are restricted to a single task: a male-female categorization. An alternative to the fixed and unidirectional CtF model, the diagnostic approach, states that differences in the task or in the stimulus drive the selection of specific SF ranges by the visual system. The diagnostic approach assumes a flexible usage of different spatial scales, even in early stages of visual processing, because they carry different diagnostic cues (Morrison & Schyns, 2001; Schyns & Oliva, 1997, 1999). Previous studies have shown that a male-female categorization relies more on low than high SFs (Deruelle & Fagot, 2005; Goffaux, Jemel, Jacques, Rossion, & Schyns, 2003). In fact, many face-encoding tasks show this preference, as low SFs has a general primacy over high SFs in face encoding (Goffaux, Gauthier, et al., 2003; Goffaux & Rossion, 2006). Nonetheless, the literature reports face categorization tasks, wherein a bias of high or middle SFs was found (X. Gao & Maurer, 2011; Schyns & Oliva, 1999; Vuilleumier et al., 2003). Therefore, it would be of great value to test the visual stimulation we implemented here, brief

dynamic sequences in CtF and FtC order, in tasks other than the male–female categorization task we used. For instance, facial identity and categorization of facial expressions tasks, as well as finer facial categorization levels, could be designed for future investigations. Indeed, it could be expected that for such finer categorization tasks, an FtC processing might be predominant. It is likely that the CtF and the diagnostic approaches might coexist and simultaneously operate in the visual system. However, it is unclear how they relate. We expect that future studies will provide a better understanding on this issue.

To summarize, the current work provides new evidence on the precedence of low over high SFs in the processing time course of human face categorization in both brain hemispheres. We used an interesting experimental design that simulates the default CtF processing and its FtC alternate mode in a divided visual field. It seems that face categorization is more stable across brain hemispheres and does not switch the order of SF integration, unlike other stimuli processing does, such as in landscapes of natural scenes that might perform FtC processing in the left hemisphere. Beyond face perception, influential models assume that the CtF strategy is the default mode of the visual system. Therefore, the current study adds evidence at the behavioral level using complex stimuli with ecological value. In future studies, it would be of great value to use this experimental design with different face-processing tasks.

References

- Amaro, E., Jr., & Barker, G. J. (2006). Study design in fMRI: Basic principles. *Brain and Cognition*, *60*, 220–232. <http://dx.doi.org/10.1016/j.bandc.2005.11.009>
- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*, 600–609. <http://dx.doi.org/10.1162/089892903321662976>
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*, 280–289. <http://dx.doi.org/10.1016/j.tics.2007.05.005>
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., . . . Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 449–454. <http://dx.doi.org/10.1073/pnas.0507062103>
- Bourne, V. J. (2006). The divided visual field paradigm: Methodological considerations. *Laterality: Asymmetries of Body, Brain and Cognition*, *11*, 373–393. <http://dx.doi.org/10.1080/13576500600633982>
- Boutet, I., Collin, C., & Faubert, J. (2003). Configurational face encoding and spatial frequency information. *Perception & Psychophysics*, *65*, 1078–1093. <http://dx.doi.org/10.3758/BF03194835>
- Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, *15*, 1411–1412. [http://dx.doi.org/10.1016/0042-6989\(75\)90200-X](http://dx.doi.org/10.1016/0042-6989(75)90200-X)
- Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, *36*(2–3), 96–107. [http://dx.doi.org/10.1016/S0165-0173\(01\)00085-6](http://dx.doi.org/10.1016/S0165-0173(01)00085-6)
- Carpenter, R. H. S. (1988). *Movements of the eyes*. London, UK: Pion.
- 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. <http://dx.doi.org/10.1037/0096-1523.30.5.975>
- Collin, C. A., Rainville, S., Watier, N., & Boutet, I. (2014). Configurational and featural discriminations use the same spatial frequencies: A model observer versus human observer analysis. *Perception*, *43*, 509–526. <http://dx.doi.org/10.1068/p7531>
- Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception*, *29*, 893–909. <http://dx.doi.org/10.1068/p2949>
- de Moraes, R., Jr., Sousa, B. M., & Fukusima, S. S. (2014). Hemispheric specialization in face recognition: From spatial frequencies to holistic/analytic cognitive processing. *Psychology & Neuroscience*, *7*, 503–511. <http://dx.doi.org/10.3922/j.pns.2014.4.09>
- Deruelle, C., & Fagot, J. (2005). Categorizing facial identities, emotions, and genders: Attention to high- and low-spatial frequencies by children and adults. *Journal of Experimental Child Psychology*, *90*, 172–184. <http://dx.doi.org/10.1016/j.jecp.2004.09.001>
- de Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545–559. [http://dx.doi.org/10.1016/0042-6989\(82\)90113-4](http://dx.doi.org/10.1016/0042-6989(82)90113-4)
- dos Santos, N. A., Andrade, S. M., & Fernandez Calvo, B. (2013). Detection of spatial frequency in

- brain-damaged patients: Influence of hemispheric asymmetries and hemineglect. *Frontiers in Human Neuroscience*, 7, 92.
- Eger, E., Schyns, P. G., & Kleinschmidt, A. (2004). Scale invariant adaptation in fusiform face-responsive regions. *NeuroImage*, 22, 232–242. <http://dx.doi.org/10.1016/j.neuroimage.2003.12.028>
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, Optics and Image Science*, 4, 2379–2394. <http://dx.doi.org/10.1364/JOSAA.4.002379>
- Gao, X., & Maurer, D. (2011). A comparison of spatial frequency tuning for the recognition of facial identity and facial expressions in adults and children. *Vision Research*, 51, 508–519. <http://dx.doi.org/10.1016/j.visres.2011.01.011>
- Gao, Z., & Bentin, S. (2011). Coarse-to-fine encoding of spatial frequency information into visual short-term memory for faces but impartial decay. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1051–1064. <http://dx.doi.org/10.1037/a0023091>
- Gauthier, I., Curby, K. M., Skudlarski, P., & Epstein, R. A. (2005). Individual differences in FFA activity suggest independent processing at different spatial scales. *Cognitive, Affective & Behavioral Neuroscience*, 5, 222–234. <http://dx.doi.org/10.3758/CABN.5.2.222>
- Goffaux, V. (2009). Spatial interactions in upright and inverted faces: Re-exploration of spatial scale influence. *Vision Research*, 49, 774–781. <http://dx.doi.org/10.1016/j.visres.2009.02.009>
- 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. [http://dx.doi.org/10.1016/S0926-6410\(03\)00056-9](http://dx.doi.org/10.1016/S0926-6410(03)00056-9)
- 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. <http://dx.doi.org/10.1068/p5370>
- Goffaux, V., Jemel, B., Jacques, C., Rossion, B., & Schyns, P. G. (2003). ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cognitive Science*, 27, 313–325. http://dx.doi.org/10.1207/s15516709cog2702_8
- Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., & Goebel, R. (2011). From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cerebral Cortex*, 21, 467–476. <http://dx.doi.org/10.1093/cercor/bhq112>
- 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. <http://dx.doi.org/10.1037/0096-1523.32.4.1023>
- Halit, H., de Haan, M., Schyns, P. G., & Johnson, M. H. (2006). Is high-spatial frequency information used in the early stages of face detection? *Brain Research*, 1117, 154–161. <http://dx.doi.org/10.1016/j.brainres.2006.07.059>
- Hegdé, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, 84, 405–439. <http://dx.doi.org/10.1016/j.pneurobio.2007.09.001>
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195, 215–243. <http://dx.doi.org/10.1113/jphysiol.1968.sp008455>
- Iidaka, T., Yamashita, K., Kashikura, K., & Yonekura, Y. (2004). Spatial frequency of visual image modulates neural responses in the temporoparietal lobe. An investigation with event-related fMRI. *Cognitive Brain Research*, 18, 196–204. <http://dx.doi.org/10.1016/j.cogbrainres.2003.10.005>
- Jacques, C., & Rossion, B. (2015). The initial representation of individual faces in the right occipitotemporal cortex is holistic: Electrophysiological evidence from the composite face illusion. *Journal of Vision*, 9, 8. <http://dx.doi.org/10.1167/9.6.8>
- Kauffmann, L., Chauvin, A., Guyader, N., & Peyrin, C. (2015). Rapid scene categorization: Role of spatial frequency order, accumulation mode and luminance contrast. *Vision Research*, 107, 49–57. <http://dx.doi.org/10.1016/j.visres.2014.11.013>
- Kauffmann, L., Ramanoël, S., & Peyrin, C. (2014). The neural bases of spatial frequency processing during scene perception. *Frontiers in Integrative Neuroscience*, 8, 37.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces – KDEF* [CD ROM]. Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet, Stockholm, Sweden.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: Freeman.
- Maurer, D., O’Craven, K. M., Le Grand, R., Mondloch, C. J., Springer, M. V., Lewis, T. L., & Grady, C. L. (2007). Neural correlates of processing facial identity based on features versus their spacing. *Neuropsychologia*, 45, 1438–1451. <http://dx.doi.org/10.1016/j.neuropsychologia.2006.11.016>
- Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin & Review*, 8, 454–469. <http://dx.doi.org/10.3758/BF03196180>
- Musel, B., Bordier, C., Dojat, M., Pichat, C., Chokron, S., Le Bas, J. F., & Peyrin, C. (2013). Reti-

- notopic and lateralized processing of spatial frequencies in human visual cortex during scene categorization. *Journal of Cognitive Neuroscience*, 25, 1315–1331. http://dx.doi.org/10.1162/jocn_a_00397
- Musel, B., Chauvin, A., Guyader, N., Chokron, S., & Peyrin, C. (2012). Is coarse-to-fine strategy sensitive to normal aging? *PLoS ONE*, 7(6), e38493. <http://dx.doi.org/10.1371/journal.pone.0038493>
- Navon, D. (1977). Forest before the trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383. [http://dx.doi.org/10.1016/0010-0285\(77\)90012-3](http://dx.doi.org/10.1016/0010-0285(77)90012-3)
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113. [http://dx.doi.org/10.1016/0028-3932\(71\)90067-4](http://dx.doi.org/10.1016/0028-3932(71)90067-4)
- Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, 34, 72–107. <http://dx.doi.org/10.1006/cogp.1997.0667>
- Parker, D. M., & Costen, N. P. (1999). One extreme or the other or perhaps the golden mean? Issues of spatial resolution in face processing. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, 18, 118–127. <http://dx.doi.org/10.1007/s12144-999-1021-3>
- Perilla-Rodríguez, L. M., de Moraes, R., Jr., & Fukusima, S. S. (2013). Lateral visual hemifield asymmetry and sex differences in recognizing low and high spatial frequency filtered faces. *Psychology & Neuroscience*, 6, 253–260. <http://dx.doi.org/10.3922/j.psns.2013.3.03>
- Peyrin, C., Schwartz, S., Seghier, M., Michel, C., Landis, T., & Vuilleumier, P. (2005). Hemispheric specialization of human inferior temporal cortex during coarse-to-fine and fine-to-coarse analysis of natural visual scenes. *NeuroImage*, 28, 464–473. <http://dx.doi.org/10.1016/j.neuroimage.2005.06.006>
- Poggio, G. F. (1972). Spatial properties of neurons in striate cortex of unanesthetized macaque monkey. *Investigative Ophthalmology*, 11, 368–377.
- Ramon, M., & Rossion, B. (2012). Hemisphere-dependent holistic processing of familiar faces. *Brain and Cognition*, 78, 7–13.
- Reinvang, I., Magnussen, S., & Greenlee, M. W. (2002). Hemispheric asymmetry in visual discrimination and memory: ERP evidence for the spatial frequency hypothesis. *Experimental Brain Research*, 144, 483–495. <http://dx.doi.org/10.1007/s00221-002-1076-y>
- Renzi, C., Schiavi, S., Carbon, C.-C., Vecchi, T., Silvano, J., & Cattaneo, Z. (2013). Processing of featural and configural aspects of faces is lateralized in dorsolateral prefrontal cortex: A TMS study. *NeuroImage*, 74, 45–51. <http://dx.doi.org/10.1016/j.neuroimage.2013.02.015>
- Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Research*, 49, 2856–2861. <http://dx.doi.org/10.1016/j.visres.2009.08.025>
- Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., De Gelder, B., & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 12, 793–802. <http://dx.doi.org/10.1162/089892900562606>
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5, 195–200. <http://dx.doi.org/10.1111/j.1467-9280.1994.tb00500.x>
- Schyns, P. G., & Oliva, A. (1997). Flexible, diagnosticity-driven, rather than fixed, perceptually determined scale selection in scene and face recognition. *Perception*, 26, 1027–1038. <http://dx.doi.org/10.1068/p261027>
- 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. [http://dx.doi.org/10.1016/S0010-0277\(98\)00069-9](http://dx.doi.org/10.1016/S0010-0277(98)00069-9)
- Sergent, J. (1982). The cerebral balance of power: Confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, 8, 253–272. <http://dx.doi.org/10.1037/0096-1523.8.2.253>
- Sergent, J. (1985). Influence of task and input factors on hemispheric involvement in face processing. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 846–861. <http://dx.doi.org/10.1037/0096-1523.11.6.846>
- Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic & Physiological Optics*, 12, 229–232. <http://dx.doi.org/10.1111/j.1475-1313.1992.tb00296.x>
- Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In N. J. Castellan, Jr. & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 199–239). Hillsdale, NJ: Erlbaum.
- Vlamings, P. H. J. M., Goffaux, V., & Kemner, C. (2009). Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency information? *Journal of Vision*, 9, 1–13. <http://dx.doi.org/10.1167/9.5.12>
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6, 624–631. <http://dx.doi.org/10.1038/nn1057>

- Watier, N. N., Collin, C. A., & Boutet, I. (2010). Spatial-frequency thresholds for configural and featural discriminations in upright and inverted faces. *Perception, 39*, 502–513. <http://dx.doi.org/10.1068/p6504>
- Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., . . . Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36*, 122–135. <http://dx.doi.org/10.1037/a0016465>
- Wu, J. H., Xu, H., Dayan, P., & Qian, N. (2010). The role of background statistics in face adaptation. *The Journal of Neuroscience, 29*, 12035–12044.
- Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special? *Vision Research, 46*, 3802–3811. <http://dx.doi.org/10.1016/j.visres.2006.06.017>

Received March 10, 2016

Revision received August 19, 2016

Accepted September 1, 2016 ■