Psychological Research (2017) 81:13–23
DOI 10.1007/s00426-015-0740-3

ORIGINAL ARTICLE

The influence of natural contour and face size on the spatial frequency tuning for identifying upright and inverted faces

Jessica Royer[1] · Verena Willenbockel[2] · Caroline Blais[1] · Frédéric Gosselin[3] · Sandra Lafortune[1] · Josiane Leclerc[1] · Daniel Fiset[1]

Received: 27 April 2015 / Accepted: 16 December 2015 / Published online: 2 January 2016

Abstract It has previously been proposed that holistic face processing is based on low spatial frequencies (SFs) whereas featural processing relies on higher SFs, a hypothesis still widespread in the face processing literature today (e.g. Peters et al. in Eur J Neurosci 37(9):1448–1457, 2013). Since upright faces are supposedly recognized through holistic processing and inverted faces, using features, it is easy to take the leap to suggest a qualitatively different SF tuning for the identification of upright and vs. inverted faces. However, two independent studies (e.g. Gaspar et al. in Vision Res 48(28):2817–2826, 2008; Willenbockel et al. in J Exp Psychol Human 36(1):122–135, 2010a) found the same SF tuning for both stimulus presentations. Since these authors used relatively small faces hiding the natural facial contour, it is possible that differences in the SF tuning for identifying upright and inverted faces were missed. The present study thus revisits the SF tuning for upright and inverted faces face identification using the SF Bubbles technique. Our results still indicate that the same SFs are involved in both upright and inverted face recognition regardless of these additional parameters (contour and size), thus contrasting with previous data obtained using different methods (e.g. Oruc and Barton in J Vis 10(12):20, 1–12, 2010). The possible reasons subtending this divergence are discussed.

✉ Daniel Fiset
daniel.fiset@uqo.ca

1 Département de Psychoéducation et de Psychologie, Université du Québec en Outaouais, C.P. 1250, Succursale Hull, Gatineau, QC J8X3X7, Canada

2 University of Victoria, Victoria, Canada

3 Université de Montréal, Montréal, Canada

Introduction

Despite the high visual similarity shared by human faces, most people are able to recognize these stimuli easily and rapidly. Interestingly, this ability can be significantly impaired by altering the face stimulus in different ways; possibly one of the most documented such manipulations is the face inversion effect (FIE; Yin, 1969). The FIE is characterized by an important drop in recognition performance (i.e. poorer accuracy and longer reaction times) when facial stimuli are rotated by 180° in the picture plane (see Rossion, 2009; Rossion & Gauthier, 2002; Valentine, 1988). It is certainly one of the most robust and replicable phenomena reported in the face processing literature.

Traditionally, the FIE has been explained by a qualitative difference between the mechanisms underlying upright and inverted face recognition (Farah, Tanaka, & Drain, 1995; Tanaka & Farah, 1993). More specifically, while upright faces would be recognized using holistic processing, the identification of inverted faces would be accomplished using piecemeal mechanisms (Rossion, 2008; Yang, Shafai, & Oruc, 2014; but see Richler, Mack, Palmeri, & Gauthier, 2011). Holistic processing is thought to rely on the processing of the face stimulus as an undecomposed whole as well as sensitivity to the relative distances between features (so-called configural processing). Piecemeal processing instead focuses on the finer details of the face, i.e. its individual features (Maurer, Le Grand, & Mondloch, 2002). In line with this idea, some have suggested that low spatial frequencies (SFs) play a key role in holistic processing, while featural or piecemeal mechanisms would mostly be subtended by high SFs (Goffaux, Hault, Michel, Vuong, & Rossion, 2005; Goffaux & Rossion, 2006; Peters, Vlamings, & Kemner, 2013). This hypothesis is supported by a handful of studies that have revealed a difference in the SF tuning for upright and inverted faces (Experiment 1 in Boutet, Collin, & Faubert, 2003; Collishaw & Hole, 2000). Nonetheless, the idea that "special" perceptual processes exist for upright face recognition has been challenged by a number of researchers (e.g. Gaspar, Bennett, & Sekuler, 2008; Gold, Mundy, & Tjan, 2012; Sekuler, Gaspar, Gold, & Bennett, 2004; Riesenhuber, Jarudi, Gilad, & Sinha, 2004; Yovel & Kanwisher, 2004; Valentine, 1988; Willenbockel et al., 2010a; Collin, Rainville, Watier, & Boutet, 2014). An alternative perspective suggests that the FIE is a mere quantitative decrease in processing efficiency for a rare and unusual stimulus, i.e. inverted faces; thus, similar perceptual processes would be solicited for recognizing faces of both orientations. In line with this idea, studies that have used a highly precise SF sampling method showed no difference in the SF tuning for upright and inverted faces (Gaspar, Sekuler, & Bennett, 2008; Willenbockel et al., 2010a). However, it is worth noting that these authors used face stimuli that seriously lacked ecological value, i.e. manipulated in a way they are rarely encountered in natural social settings. Indeed, the stimuli presented to the observers were shaped as ellipses (the natural contours of the faces were hidden) and were also quite small (under 6° of visual angle).

Recent evidence suggests that it is crucial to maximize the ecological value of the face stimuli in order to adequately measure face-specific perceptual processes (McKone, 2009; Oruc & Barton, 2010; Yang et al., 2014; Ross & Gauthier, 2015). Considering this, the present study revisits the issue of SF tuning for upright and inverted faces while taking into account two additional ecologically relevant parameters that may influence SF tuning: the presence of natural facial contour and face size. Given that external facial features such as natural facial contours do contain information useful for face identification (Ellis, Shepherd, & Davies, 1979; Chan & Ryan, 2012; Hills, Romano, Davies-Thompson, & Barton, 2014; but see Butler, Blais, Gosselin, Bub, & Fiset, 2010), it is possible that real-world differences between upright and inverted face SF processing were missed in previous studies using smaller face stimuli. Thus, we investigate in Experiment 1 the, to the best of our knowledge, unexplored issue of the role of natural facial contour in the SF tuning of upright and inverted faces. As for the role of stimulus size, some have suggested that faces must exceed a certain size criterion (approximately 5°–7° of visual angle; Oruc & Barton, 2010; Yang et al., 2014) for natural (i.e. configural) face processing mechanisms to be set into motion. Previous results by Oruc and Barton (2010) showed that SF tuning for upright and inverted faces differs qualitatively as a function of size: Indeed, their data indicate that while the critical SFs in letter, novel shape and inverted face identification are scale dependant, upright face identification displays a scale-invariant pattern at widths over approximately 5° of visual angle. Thus, a divergence in the SF tuning for both orientations as a function of stimulus size would be apparent starting at this face width. However, they do not present a direct statistical comparison of the SF tuning for both orientation conditions at these larger sizes. Also, the SF sampling procedure used in this study (critical-band masking) measures SF tuning by pre-selecting specific frequency bands. In Experiment 2, we directly investigate whether the results obtained by Oruc and Barton (2010) suggesting a qualitative difference between both orientations at larger faces sizes (but optimally at approximately 9 degrees of visual angle; Yang et al., 2014) can be obtained using a SF sampling method that randomly selects multiple SFs on a trial-by-trial basis in an unbiased manner. With this second Experiment, we also verify if the conclusions presented in previous papers supporting only a quantitative difference in SF tuning between both orientations (Gaspar et al., 2008; Willenbockel et al., 2010a) still stands for larger face stimuli.

Experiment 1

The first objective of this study was to examine the SF tuning for upright and inverted face identification in two different experimental conditions, i.e. when the natural contours of the face were preserved and when they were hidden (i.e. cropped to an elliptical shape). This objective was addressed using SF Bubbles, a variant of the Bubbles method (Gosselin & Schyns, 2001; Tadros, Dupuis-Roy, Fiset, Arguin & Gosselin, 2013; Royer, Blais, Gosselin, Duncan, & Fiset, 2015; Thurman, & Grossman, 2011; Willenbockel et al., 2010a; Willenbockel, Lepore, Bacon, & Gosselin, 2013; Willenbockel, Lepore, Nguyen, Bouthillier, & Gosselin, 2012). The basic idea behind SF Bubbles is that by randomly sampling certain SFs on a trial-by-trial basis, we will be able to pinpoint, after many trials, which SFs are significantly correlated with accuracy for identifying upright and inverted faces.

Method

Participants

Sixteen participants took part in Experiment 1. All participants had normal or corrected-to-normal vision, and were naïve to the purpose of the experiment. The Université du Québec en Outaouais's Research Ethics Committee approved the study, and all participants provided written consent.

Apparatus

The experiment ran on a 2.5-GHz Macintosh computer (model: dual 2.5 GHz PowerPC G5). Stimuli were displayed on a Sony Trinitron Multiscan G420 monitor measuring 45.4 cm in diagonal, with a resolution of 1024 × 768 pixels and a refresh rate of 100 Hz. The monitor was calibrated to allow a linear manipulation of luminance. The adjusted lookup table contained 167 luminance levels, ranging from 0.97 cd/m^2 to 138.67 cd/m^2. The experimental programs were developed in Matlab (Natick, MA) using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The viewing distance was maintained at 51 cm using a chinrest.

Stimuli

We created a database composed of 20 faces (10 women). There were two versions of each face stimulus: one where the natural contour of the face was preserved, and one on which we applied an elliptical aperture hiding the natural facial contour. We then pseudo-randomly assigned each identity to one of two sets of equal size (see Fig. 1). Eight participants saw the faces from Set 1 with contour and the faces from Set 2 without contour, whereas the eight other participants saw the faces from Set 1 without contours and faces from Set 2 with contour. Face width was 5 degrees of visual angle (equivalent to a face seen from approximately 1.6 m), and image resolution was 256 × 256 pixels. Faces within each stimulus set were equated in mean luminance, contrast, and energy at each SF using SHINE (Willenbockel et al., 2010b).

Learning phase

Participants learned to associate the faces with common French Canadian names (e.g. Caroline, Cynthia, Vincent…) from printed grayscale pictures displayed along with these names. Each of the numerals (0–9) on a regular computer keyboard were associated with a particular face name. When the participants were confident that they could identify all 20 faces, the practice session began. The stimuli from each of the four conditions (Upright/Contour, Upright/Ellipse, Inverted/Contour, or Inverted/Ellipse) were presented in separate 100-trial blocks. Each trial began with a central fixation cross on the screen for 100 ms, followed by a face belonging to one of the four conditions. The face stimulus remained onscreen until the observer responded with a key press. The first part of the learning session was completed when performance was above 95 % correct for each of the four conditions, consecutively. Finally, the participants completed twelve additional practice blocks (three blocks of each condition) in which white Gaussian noise was added to the broadband faces. This second learning phase was conducted in preparation for the experimental blocks. The objective of this second phase was to familiarize the participants with stimuli that resembled those used in the experimental phase. Indeed, during the SF Bubbles task, the faces were sampled in the SF domain and white Gaussian noise was added to the sampled stimulus in order to control participants' performance (see "Experimental phase" section below). The signal-to-noise ratio (SNR) started at 1.5 for all subjects. The 256 × 256 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 in increments of 0.05 by the experimenter in order to maintain accuracy minimally at 90 %. The starting SNR and the 0.05 increment were chosen based on the results of a pilot study conducted prior

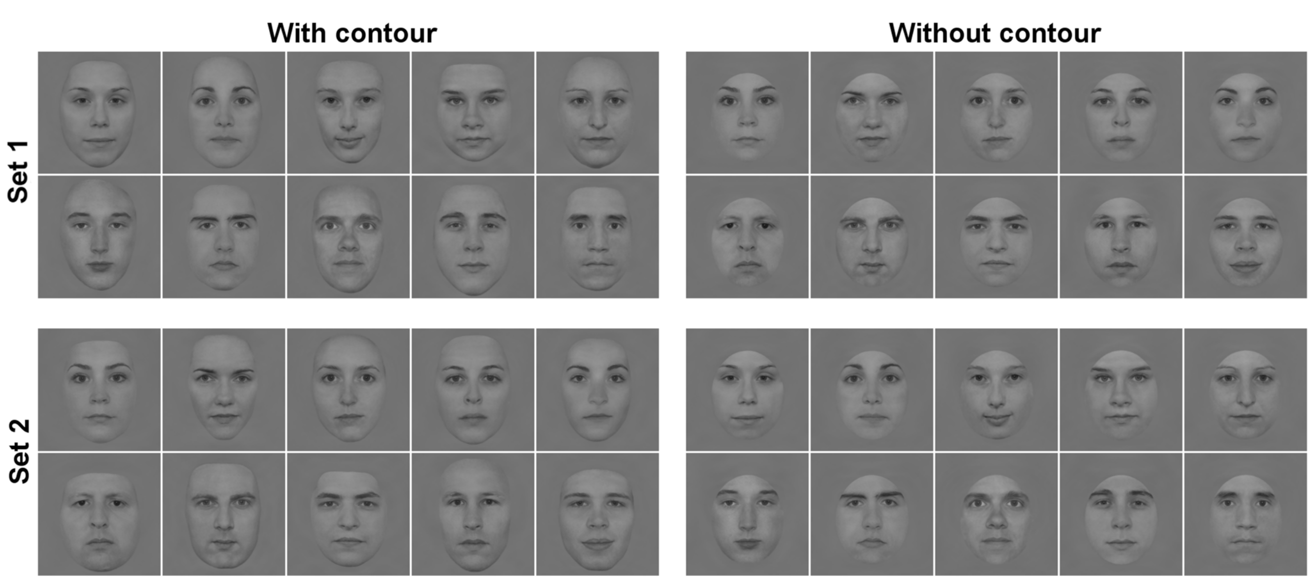

Fig. 1 Stimuli used in Experiments 1 and 2

to this experiment (see Willenbockel et al., 2010a, b, for a similar design).

Experimental phase

The SF Bubbles method was created using the same underlying logic as the Bubbles method (Gosselin & Schyns, 2001). The basic idea behind Bubbles is that by randomly sampling specific visual information on a trial-by-trial basis, we will be able to precisely determine, after many trials, what information is significantly correlated with observers' performance. In our case, SF Bubbles revealed which SFs correlated with accurate identification of upright and inverted faces, as well as the influence of contour information on the SF tuning for each orientation condition. The next paragraph presents a detailed account of the technique by which the facial stimuli were sampled using SF Bubbles.

At any given trial, the SF information of a stimulus was sampled randomly as illustrated in Fig. 2; this procedure is identical to what is described in Willenbockel et al. (2010a). First, the base stimulus was padded with a uniform gray background of twice the stimulus' size in order to minimize edge artifacts in the SF domain. Second, the padded stimulus was fast Fourier transformed (FFT) using functions from the Image Processing Toolbox for MATLAB. Steps 3a to 3d illustrate the construction of the random SF filter. The SF filter was built by (3a) creating a random binary vector of $2wk$ elements (35 ones among 10,205 zeros, 35 being the number of bubbles), where w is equal to the stimulus' width, and k is a constant, set here at 20, that determines the smoothness of the sampling. As shown in (3b), a smooth filter was constructed by convoluting the binary vector with a Gaussian kernel (a "SF Bubble"; $\sigma = 1.8$), resulting in a random sampling vector. This smoothed vector was then subjected to a logarithmic transformation in step (3c) in order to fit the human visual system's SF sensitivity (De Valois & De Valois, 1990). This transformation renders the bubbles that sample lower SFs much narrower than those that sample higher SFs. The final step (3D) in the creation of the SF filter was the rotation of the log-transformed vector about its origin to obtain a 2D filter. Filtering was carried out in step 4 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. The experimental stimulus was finally constructed by cropping the central 256 × 256 pixel region of the filtered image.

During the experimental phase, participants completed 1200 trials of each condition, for a total of 4800 trials. The task design was the same as in the practice phase, i.e. a 10-choice identification task. The only differences between the practice phase and the experimental phase were that (a) the SFs of the base stimuli were randomly sampled using SF Bubbles, and (b) performance in the upright blocks was maintained between 75 and 85 % correct by adjusting the quantity of additive noise on a block-per-block basis (see Willenbockel et al., 2010a, for a similar design). The same amount of noise required to keep individual performance within this interval in the upright blocks was added to the inverted stimuli. The amount of noise was manipulated independently for each upright contour condition. The SNR started at 2 for all subjects. The value of c was increased or decreased on a block-by-block basis by the experimenter (in increments of 0.02). The starting SNR and the 0.02 increment were chosen based on the results of a pilot study conducted prior to this experiment.

Results and discussion

The first 100 experimental trials of each condition were considered practice trials for the Bubbles method and were not taken into account in the analyses. We first verified that our procedure to control accuracy for upright faces in both contour conditions worked as planned, which was confirmed by a t test ($M_{contour} = 76.55$ %, $SD_{contour} = 1.92$ %; $M_{ellipse} = 76.31$ %, $SD_{ellipse} = 1.86$ %; $t(15) = .49$, ns). A two-way repeated measures ANOVA (orientation conditions × contour conditions; bonferroni-corrected) on the accuracy data revealed a significant effect of orientation ($M_{upright} = 76.43$ %, $SD_{upright} = 1.62$ %; $M_{inverted} = 51.18$ %, $SD_{inverted} = 5.98$ %; $F(1,15) = 364.98$, $p < .01$; partial eta squared [η_p^2] = .96), of contour ($M_{contour} = 65.24$ %, $SD_{contour} = 3.88$ %; $M_{ellipse} = 62.37$ %, $SD_{ellipse} = 4.11$ %; $F(1,15) = 8.74$, $p = .01$; $\eta_p^2 = .37$), and, most importantly, a significant interaction between orientation condition and contour condition ($F(1,15) = 8.85$, $p = .01$; $\eta_p^2 = .37$). This means that even if performance was strictly controlled in both contour conditions for upright faces, the negative impact of inversion on recognition performance was stronger for faces presented through an elliptical aperture compared to faces for which the natural contour was preserved. In other words, including the natural facial contours benefited recognition performance for inverted faces. Similarly, the participants managed to attain the required accuracy criterion (for upright faces) with less signal when the contours were present than when they were hidden, as we can see by comparing the values of c for each contour condition ($M_{contour} = .51$, $SD_{contour} = .03$; $M_{ellipse} = .61$, $SD_{ellipse} = .05$; $t(15) = -7.82$, $p < .01$; $d = -2.49$). Thus, it seems that including natural facial contour benefited performance in both orientation conditions.

To find out which SFs drove the observers' correct and incorrect responses, we performed a multiple linear

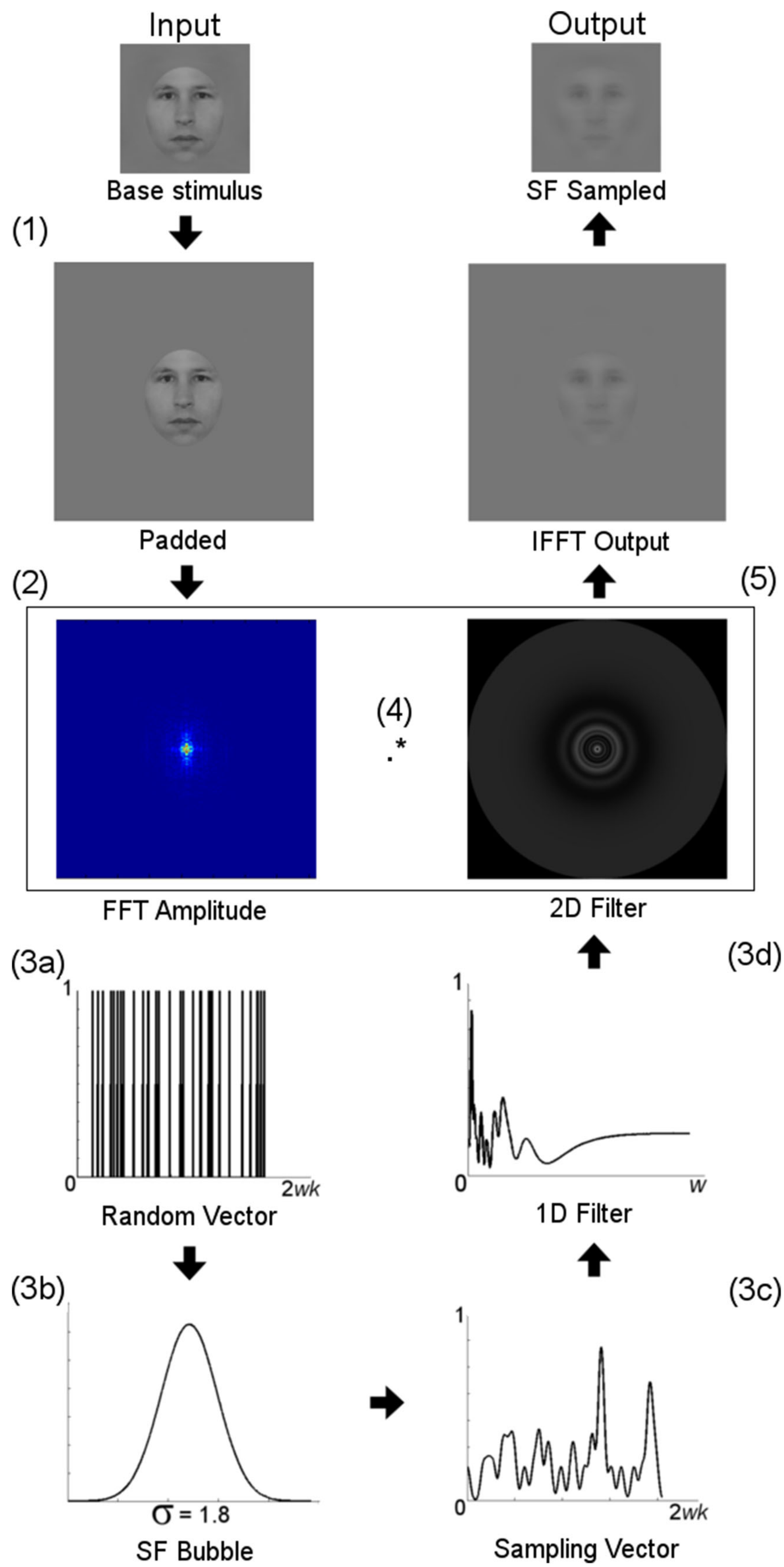

Fig. 2 Creation of a Bubblized stimulus using the SF Bubbles technique (see text for details). Note that we zoomed the image in step 2 towards its center in order to better illustrate the stimulus' complex FFT amplitudes

regression on the aforementioned random binary vector (i.e. the sampled SFs) and the observers' accuracy. The vector of regression coefficients—referred to as a classification vector—was then transformed into Z scores for each observer using the mean and standard deviation of each individual observer. The group classification vector was computed by summing the classification vectors of all observers and by dividing the resulting vector by the square root of the number of observers (i.e. 16). A pixel test was used to determine a statistical threshold (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). As in Tadros et al. (2013), we measured the SF peaks by submitting the classification vector to a 50 % area spatial frequency measure (ASFM; analogous to a 50 % area latency measure commonly used in electroencephalography analysis). This 50 % ASFM is a better measure of the central tendency than absolute peaks since it is less sensitive to the shape of the tuning curve, which was probably slightly distorted by our sampling procedure.

To reveal the SF ranges that led to accurate face identification in the upright and the inverted conditions, we performed multiple linear regressions on the sampling vectors and accuracy per condition per observer. As illustrated in Fig. 3, all conditions' group classification vectors showed SF bands that were significantly correlated with face recognition accuracy. Figure 3 compares the SF tuning for both orientation conditions independently for each contour condition. When the face stimuli were shown through an elliptical aperture (Fig. 3, top), the upright group classification vector (full line) showed a significant SF band of 2.7 octaves peaking at 9.7 cpf ($Z_{max} = 16.3$; $p < .05$; Full width half maximum [FWHM] = 5.89; $Z_{criterion} = 3.30$, statistical threshold illustrated by the dashdotted line in Fig. 3, two-tailed), and the inverted group classification vector (dashed line) showed a significant 2.8 octave wide SF band peaking at 9.6 cpf ($Z_{max} = 12.7$). The difference classification vector did not reach statistical significance (dotted line). In the natural contour condition (Fig. 3, bottom), the upright classification vector (full line) showed a significant SF band of 3.1 octaves peaking at 7.3 cpf ($Z_{max} = 17.1$). Similarly, in the inverted condition (dashed line), a SF range of 2.9 octaves peaking at 6.7 cpf ($Z_{max} = 15.4$) reached significance. As in the elliptical condition, the difference classification vector comparing both orientation conditions did not reach statistical significance (dotted line). Thus, according to our results, even when adding ecological value to the face stimuli by including natural facial contour information, the SF tuning for upright and inverted faces does not differ in the way predicted by the qualitative framework of the FIE.

The effect of including natural contour information seems to be the same for both orientation conditions. The difference classification vector comparing both contour conditions for upright faces showed a 1.2 octave wide significant SF band peaking at 3.9 cpf ($Z_{max} = 4.9$). For inverted faces, the difference classification vector showed two significant SF bands: one 2.7 octaves wide peaking at 3.6 cpf ($Z_{max} = 4.7$), and one 2.5 octaves wide reaching its minimum at 42.2 cpf ($Z_{min} = -4.2$). Thus, including the natural contour of the face significantly shifted the observers' SF tuning towards lower SFs for both upright and inverted face recognition, which is congruent with our accuracy and SNR data. This means that the observers were able to adapt their visual strategy to the additional information conveyed by natural facial contours, and used this information in a way that benefited their recognition accuracy. It is interesting to note that, at least for faces of the size used in this Experiment, a manipulation that increases reliance on low SFs (i.e. including natural contour) decreases the FIE at the performance level.

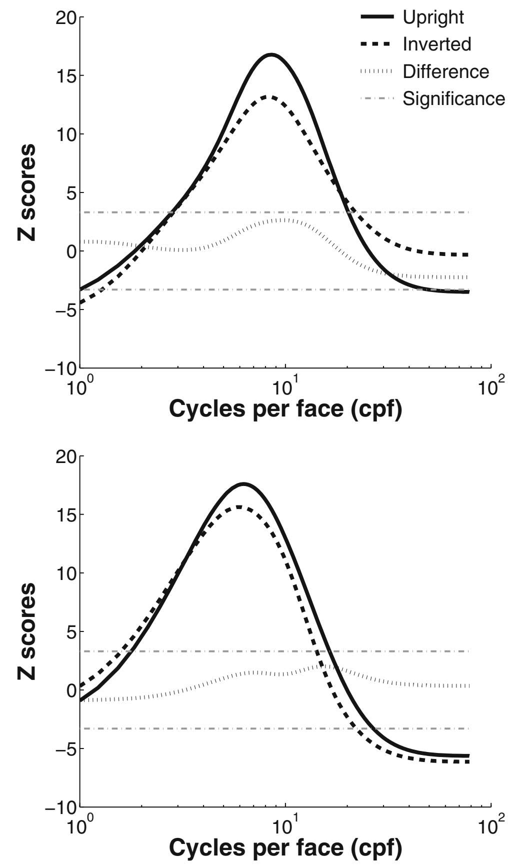

Fig. 3 SF tuning for upright and inverted faces when the natural facial contours were absent (*top*) and present (*bottom*). Neither difference classification vector reached significance (the statistical threshold is defined by the *dash-dotted line*)

Despite our attempt to increase the ecological value of our face stimuli by including natural facial contour, the absence of a significant difference in SF tuning between upright and inverted faces could be explained by our use of relatively small stimuli (as in Willenbockel et al., 2010a, b, and Gaspar et al., 2008, that used faces of approximately 6° and 2.3° of visual angle, respectively). Indeed, the absence of a difference between both orientations is in line with Oruc and Barton's (2010) findings: these authors suggest that a qualitative difference in the SF tuning for upright and inverted faces is only present at larger stimulus sizes, i.e. those typically encountered by the visual system in an ecological setting (i.e. at a casual discussion distance). Recently, Yang et al. (2014) proposed that face width must be at least 7° of visual angle (but ideally 8–10 degrees, which is equivalent to seeing a real face from about 0.8–1 m) for face-specific identification mechanisms to be recruited. If the face stimuli fail to meet this size criterion, the difference between upright and inverted face processing efficiency would be only quantitative, and not qualitative. In light of these findings, we conducted a second experiment where we investigate whether Oruc and Barton's (2010) findings generalize to a different task using different facial stimuli and a different SF sampling method, i.e. SF bubbles.

Experiment 2

Experiment 2 sought to investigate whether a significant difference in SF tuning between both orientation conditions can be found when using larger face stimuli, and if this potential difference would be in line with the qualitative framework's predictions.

Method

Participants

Sixteen participants completed the task. None of these participants had taken part in Experiment 1. All participants had normal or corrected-to-normal vision, and were naïve to the purpose of the experiment. The Université du Québec en Outaouais's Research Ethics Committee approved the study, and all participants provided written consent.

Apparatus

The experiment was conducted on MacPro QuadCore computers. Stimuli were displayed on a 22-inch 120 Hz Samsung LCD monitor. The monitor's resolution was set to 1680 × 1050 pixels. The viewing distance was maintained at 50 cm using a chinrest.

Stimuli

As in Experiment 1, participants had to identify a total of 20 faces (10 in each contour condition); half of the participants were tested with the faces from Set 1 with contours and the faces from Set 2 without contours, whereas the others saw the faces from Set 1 without contours and faces from Set 2 with contours. Face width was 9 degrees of visual angle, and image resolution was 512 × 512 pixels.

Procedure

The experimental design was identical to Experiment 1, i.e. consisted of a learning phase (noise-free trials followed by added noise trials) and an experimental phase (with SF bubbles). In Experiment 2, the amount of noise added to the stimuli in the learning phase and during the bubbles task was manipulated on a trial-by-trial basis using QUEST (Watson & Pelli, 1983), instead of being adjusted by the experimenter on a block-by-block basis. The observers' accuracy was maintained at 75 %. The amount of noise was only adjusted by QUEST during upright trials; the amount of noise in the inverted conditions was set to the amount of noise suggested by QUEST at the last trial of the upright condition. Our participants were asked to complete 1100 trials of each condition, for a total of 4400 trials. The data from Experiment 2 were subjected to the same analyses as previously described in Experiment 1.

Results and discussion

As in Experiment 1, the first 100 trials of the experimental phase were considered practice trials and were not taken into account in the analyses. We first verified that our procedure to control accuracy for upright faces in both contour conditions worked as planned, which was confirmed by a *t* test ($M_{contour} = 75.42$ %, $SD_{contour} = 6.89$ %; $M_{ellipse} = 76.04$ %, $SD_{ellipse} = 6.36$ %; $t(15) = -.55$, ns). Unsurprisingly, upright face recognition accuracy was extremely similar in Experiments 1 and 2. We then conducted a two-way repeated measures ANOVA (orientation conditions × contour conditions; bonferroni-corrected) on the participants' accuracy which revealed a significant effect of orientation ($M_{upright} = 75.73$ %, $SD_{upright} = 6.24$ %; $M_{inverted} = 52.04$ %, $SD_{inverted} = 7.64$ %; $F(1,15) = 306.91$, $p < .01$; $\eta_p^2 = .95$), but not of contour ($M_{contour} = 63.91$ %, $SD_{contour} = 6.31$ %; $M_{ellipse} = 63.87$ %, $SD_{ellipse} = 6.99$ %; $F(1,15) = .00$, ns), and no interaction between orientation and contour condition ($F(1,15) = .64$, ns). Similarly, no significant difference was revealed when comparing the required signal for the with contour and without contour conditions ($M_{contour} = .82$, $SD_{contour} = .07$; $M_{ellipse} = .82$, $SD_{ellipse} = .05$; $t(15) = -.57$, ns).

All conditions' group classification vector showed SF bands significantly correlated with face recognition. When

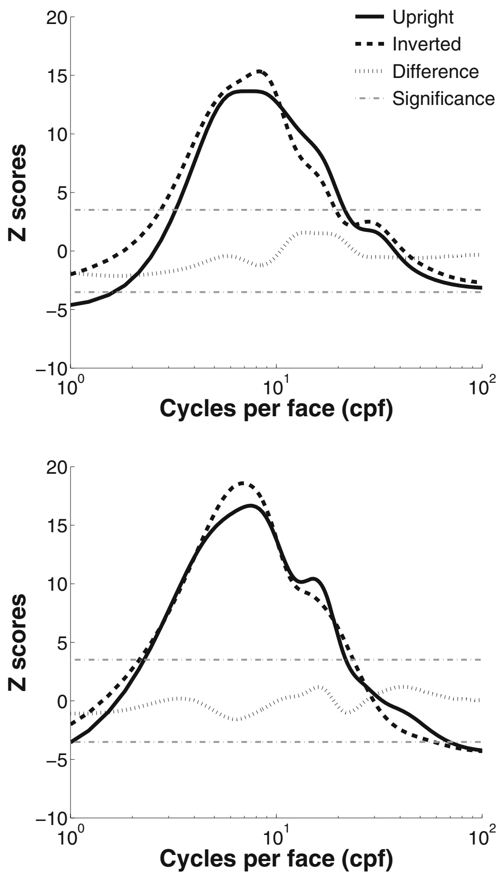

Fig. 4 SF tuning for upright and inverted faces when the natural facial contours were absent (*top*) and present (*bottom*). Neither difference classification vector reached significance (the statistical threshold is defined by the *dash-dotted line*)

the face stimuli were shown through an elliptical aperture (Fig. 4, top), the upright group classification vector (full line) showed a significant SF band of 2.7 octaves peaking at 9.8 cpf ($Z_{max} = 12.9$, $p < .05$, $FWHM = 5.89$, $Z_{criterion} = 3.51$; statistical threshold illustrated by the dash-dotted line in Fig. 4, two-tailed), and the inverted group classification vector (dashed line) showed a significant 2.8 octave wide SF band peaking at 8.5 cpf ($Z_{max} = 15.3$). The difference classification vector did not reach statistical significance. Furthermore, when the natural contours of the face were included in the stimuli (Fig. 4, bottom), the upright classification vector (full line) showed a significant SF band of 3.1 octaves peaking at 9.1 cpf ($Z_{max} = 15.4$). Similarly, in the inverted condition (dashed line), a SF range of 3.3 octaves peaking at 8.7 cpf ($Z_{max} = 16.9$) reached significance. Again, the difference classification vector comparing both orientation conditions did not reach statistical significance. Thus, even when using larger faces, our findings do not support the qualitative framework's prediction concerning SF tuning differences between upright and inverted faces.

Furthermore, the difference classification vector comparing both contour conditions for upright faces shows a significant .59 octave wide difference peaking at 2.95 cpf ($Z_{max} = 3.89$). However, the difference classification vector comparing both contour conditions for inverted faces did not reach statistical significance. Thus, the visual information contained in natural facial contours (i.e. lower SFs) was useful in upright, but not inverted, face recognition. This contrasts with the results of Experiment 1, where natural facial contours induced a significant shift towards lower SFs for both orientations. This suggests that when confronted with larger, more ecological face stimuli, the observers' use of SF information seems more flexible with upright than inverted faces. This greater use of lower SFs for upright faces is nonetheless surprising seeing as the presence of natural facial contour does not benefit performance or lower the SNR as was observed in Experiment 1 with smaller face stimuli. It thus seems that the observers can use contour information, but this different visual strategy does not significantly alter the quality of their performance. The fact that we observe a change in SF tuning but not in performance suggests that contour information can be useful in upright face recognition. However, this information would be used independently from other visual cues, meaning that the participants did not use both the contour and other facial cues during the same trials. In the end, this resulted in natural facial contour not contributing any advantage to response accuracy.

General discussion

The FIE is traditionally explained by the use of qualitatively distinct perceptual mechanisms for the identification of inverted and upright faces (e.g. Goffaux et al., 2005; Farah et al., 1995; Tanaka & Farah, 1993; Rossion, 2008). Those that have directly tested the outcome of this hypothesis on the FIE did not reveal any significant SF tuning difference between both orientation conditions (Gaspar et al., 2008; Willenbockel et al., 2010a). However, these authors used rather small facial stimuli (between 5 and 6° of visual angle or less) that were presented through an elliptical aperture, thus hiding the natural contour of the face. In light of recent evidence suggesting that both of these ecologically relevant parameters seem to influence natural face recognition mechanisms (e.g. Hills et al., 2014; Oruc & Barton, 2010; Yang et al., 2014), the present paper aimed at revisiting the SFs diagnostic for upright and inverted faces using stimuli that take these parameters into account.

In Experiment 1, we compared the SF tuning for upright and inverted smaller faces (5° of visual angle) when these stimuli were presented through an elliptical aperture versus when the natural facial contours were preserved. Experiment 2 compared the same conditions as Experiment 1 using larger face stimuli (9° of visual angle). Both Experiments replicated the FIE at the performance level, but did not reveal any significant difference in SF tuning between both orientation conditions, regardless of stimulus size and the presence or absence of natural facial contour. In fact, the SFs that were the most correlated with accuracy for both upright and inverted faces are in line with what has been demonstrated in past studies (i.e. intermediate SFs; Oruc & Barton, 2010; Gaspar et al., 2008; Willenbockel et al., 2010a, Peli, Lee, Trempe, & Buzney, 1994, Näsänen, 1999; Gold, Bennett, & Sekuler, 1999; Collin et al., 2014). Thus, our results bring further support to the hypothesis that the FIE is due to quantitative differences in the efficiency with which information from the same SF band is used in both orientations (e.g. Gold et al., 2012; Riesenhuber et al., 2004; Sekuler et al., 2004; Richler et al., 2011). We show that while face inversion does not seem to induce qualitative changes in SF tuning, including natural facial contour does significantly increase the diagnosticity of low SFs. Interestingly, this effect of contour differed depending on the size of the face stimuli: While contour information affected SF tuning in the same way for both upright and inverted smaller faces, the shift towards lower SFs was only observed for upright, but not inverted, larger faces.

Recent studies have suggested that the SF tuning for upright and inverted faces is qualitatively different specifically with faces of a size typically encountered in social contexts (i.e. over approximately 6 degrees of visual angle; Oruc & Barton, 2010; Yang et al., 2014). These authors explain the difference in SF tuning between upright and inverted faces by a flattening of the SF tuning curve for upright faces past a width of approximately 4.7° of visual angle, while the SF tuning curve for inverted faces continues to increase as a function of face size. Based on Oruc and Barton's results and the size of our stimuli, we expected no size effect on the peak of the spatial frequency tuning with upright faces, but a shift in the position of the peak in favor of higher spatial frequencies with inverted faces. We replicated Oruc & Barton's findings with upright faces, but not with inverted faces. Indeed, with inverted faces, the peak remained at a similar spatial frequency thus suggesting that the same perceptual strategies can be applied with both orientations. Note that we only compared our results from the ellipse condition to those of Oruc and Barton (2010), since their stimuli were presented through an ellipse.

Many methodological differences between our study and Oruc and Barton's (2010) could explain the different results obtained with inverted faces. For example, we used a filtering method whereas they used critical-band masking. Indeed, in the field of face recognition, these two methods have sometimes produced contrasting results (Gold, Bennett & Sekuler, 1999; but see Willenbockel et al., 2010a, and Gaspar et al., 2008 for converging results using both methods). Another difference concerns the accuracy criterion for threshold measurement. While Oruc and Barton (2010) measured recognition threshold in both orientations at quite high and identical accuracy level, i.e. 82 % correct, we decided to equalize the amount of noise in both orientations by finding the amount of noise leading to an accuracy rate of 75 % in the upright condition, and using this same level of noise in the inverted condition. This led to an accuracy drop with inverted faces (52 % on average). Yet another factor that could have contributed to the divergent findings is the number of response alternatives. We used 10 identities, while Oruc and Barton used five.

.Furthermore, one could argue that a lack of statistical power could account for our failure to reveal a significant difference between both orientation conditions. However, this hypothesis seems, to us, unlikely: Firstly, our results show, overall, near zero difference between upright and inverted faces in lower SFs. It is difficult to believe that any amount of additional participants could change this null effect to eventually reveal significant findings. Also, when combining both contour conditions separately for each experiment, the difference classification vectors between upright and inverted faces still does not reach statistical significance, despite maximum Z scores reaching approximately 25 in each orientation condition. Lastly, it is important to note that we do succeed in revealing significant differences in the SF tuning between both contour conditions, which further argues against a lack of statistical power.

Conclusion

In summary, our data do not directly support the existence of a qualitative difference in the SF tuning for upright and inverted faces. In fact, based on our results, it seems that even if such a difference does exist, it appears specific to particular experimental properties that are not mandatory to reveal a FIE at the performance level, and of negligible magnitude in respect to SF tuning. Nonetheless, the few discrepancies revealed between our data and previous results support the importance of continuing to investigate the role of SFs and stimulus size in the FIE using various SF sampling methods, as well as diverse types of face recognition and control tasks. Although our results indicate that the use of SF information does not seem to be a major factor in explaining the FIE, recent studies investigating the use of orientation information revealed considerable differences between the orientation tuning for upright and inverted faces. A handful of studies have demonstrated the important role played by horizontal information in face identification (Dakin & Watt, 2009; Goffaux & Dakin, 2010; Goffaux, van Zon, & Schiltz, 2011; Pachai, Sekuler, & Bennett, 2013). Interestingly, Goffaux and collaborators showed that this tuning to horizontal information in upright faces was most pronounced in intermediate SF bands, followed by high SFs, and absent in low SFs. They also showed that face inversion disrupted the processing of horizontal information in middle and high SF bands (Goffaux et al., 2011). Nonetheless, the SFs that they identified as critical for face recognition are not in line with the qualitative framework's predictions as presented in the introduction, i.e. that holistic processing (upright face recognition) is subtended by low SFs and piecemeal processing by high SFs (inverted face recognition; e.g. Dobkins & Harms, 2014). The manipulation of orientations by themselves as well as in combination with SFs thus seems to be a promising avenue to better understand the visual mechanisms subtending the FIE.

Acknowledgments This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Daniel Fiset, Caroline Blais, and Frédéric Gosselin, by a scholarship from the Fonds de Recherche du Québec Nature et Technologies and from NSERC to Jessica Royer and a summer scholarship from NSERC to Sandra Lafortune and Josiane Leclerc.

References

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception and Psychophysics, 65*(7), 1078–1093.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics, 72*(6), 1444–1449.

Chan, J. P., & Ryan, J. D. (2012). Holistic representations of internal and external face features are used to support recognition. *Frontiers in Psychology, 3*, 87.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision, 5*(9), 659–667.

Collin, C. A., Rainville, S., Watier, N., & Boutet, I. (2014). Configural and featural discriminations use the same spatial frequencies: A model observer versus human observer analysis. *Perception, 43*(6), 509–526.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*(8), 893–909.

Dakin, S. C., & Watt, R. J. (2009). Biological "bar codes" in human faces. *Journal of Vision*, *9*(4):2, 1–10.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial vision*. New York: Oxford University Press.

Dobkins, K. R., & Harms, R. (2014). The face inversion effect in infants is driven by high, and not low, spatial frequencies. *Journal of Vision*, *14*(1), 1, 1–17.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception, 8*(4), 431–439.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 628–634.

Gaspar, C., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research, 48*(28), 2817–2826.

Goffaux, V., & Dakin, S. C. (2010). Horizontal information drives the behavioral signatures of face processing. *Frontiers in Psychology, 1*, 143.

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*(1), 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*(4), 1023–1039.

Goffaux, V., van Zon, J., & Schiltz, C. (2011). The horizontal tuning of face perception relies on the processing of intermediate and high spatial frequencies. *Journal of Vision*, *11*(10), 1, 1–9.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39*(21), 3537–3560.

Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The perception of a face is no more than the sum of its parts. *Psychological Science, 23*(4), 427–434.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*(17), 2261–2271.

Hills, C., Romano, K., Davies-Thompson, J., & Barton, J. J. (2014). An adaptation study of internal and external features in facial representations. *Vision Research, 100*, 18–28.

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255–260.

McKone, E. (2009). Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. *Vision Research, 49*(2), 268–283.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39*(23), 3824–3833.

Oruç, İ., & Barton, J. J. (2010). Critical frequencies in the perception of letters, faces, and novel shapes: Evidence for limited scale invariance for faces. *Journal of Vision*, *10*(12), 20, 1–12.

Pachai, M. V., Sekuler, A. B., & Bennett, P. J. (2013). Sensitivity to information conveyed by horizontal contours is correlated with face identification accuracy. *Frontiers in psychology*, *4*:74.

Peli, E., Lee, E., Trempe, C. L., & Buzney, S. (1994). Image enhancement for the visually impaired: The effects of enhancement on face recognition. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 11*(7), 1929–1939.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Peters, J. C., Vlamings, P., & Kemner, C. (2013). Neural processing of high and low spatial frequency information in faces changes across development: Qualitative changes in face processing during adolescence. *European Journal of Neuroscience, 37*(9), 1448–1457.

Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research, 51*(3), 333–342.

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 of London. Series B: Biological Sciences, 271*(Suppl 6), S448–S450.

Ross, D., & Gauthier, I. (2015). Holistic processing in the composite task depends on face size. *Visual Cognition, 23*(5), 533–545.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*(2), 274–289.

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica, 132*(3), 300–312.

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

Royer, J., Blais, C., Gosselin, F., Duncan, J., & Fiset, D. (2015). When less is more: Impact of face processing ability on recognition of visually degraded faces. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1179-1183.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14*(5), 391–396.

Tadros, K., Dupuis-Roy, N., Fiset, D., Arguin, M., & Gosselin (2013). Reading laterally: The cerebral hemispheric use of spatial frequencies in visual word recognition. *Journal of Vision, 13*(1), 4, 1–12.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology, 46*(2), 225–245.

Thurman, S. M., & Grossman, E. D. (2011). Diagnostic spatial frequencies and human efficiency for discriminating actions. *Attention, Perception, & Psychophysics, 73*(2), 572–580.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*(4), 471–491.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics, 33*(2), 113–120.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., … Gosselin, F. (2010a). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36*(1), 122–135.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010b). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods, 42*(3), 671–684.

Willenbockel, V., Lepore, F., Nguyen, D. K., Bouthillier, A., & Gosselin, F. (2012). Spatial frequency tuning during the conscious and non-conscious perception of emotional facial expressions—an intracranial ERP study. *Frontiers in Psychology, 3*, 237.

Willenbockel, V., Lepore, F., Bacon, B. A., & Gosselin, F. (2013). The informational correlates of conscious and nonconscious face-gender perception. *Journal of Vision*, *13*(2), 10, 1–14.

Yang, N., Shafai, F., & Oruc, I. (2014). Size determines whether specialized expert processes are engaged for recognition of faces. *Journal of Vision*, *14*(8), 17, 1–12.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141–145.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron, 44*(5), 889–898.