Vision Research 51 (2011) 508–519

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

A comparison of spatial frequency tuning for the recognition of facial identity and facial expressions in adults and children

Xiaoqing Gao, Daphne Maurer *

McMaster University, Canada

ARTICLE INFO

Article history:
Received 13 March 2010
Received in revised form 23 January 2011
Available online 28 January 2011

Keywords:
Facial identity
Facial expression
Spatial frequency
Development
Efficiency

ABSTRACT

We measured contrast thresholds for the identification of faces and facial expressions as a function of the center spatial frequency of narrow-band additive noise. In adults, masking of mid spatial frequencies (11–16 c/fw) caused the largest elevation in contrast threshold (Experiment 1). Ideal observer analysis revealed that adults were equally sensitive to available information at low and mid spatial frequencies, both of which they used more efficiently than high spatial frequencies. The drop-off of sensitivity at high spatial frequencies began at a lower spatial frequency for recognizing facial identity than for recognizing facial expression. As a result, the critical band was higher for expression than for identity. The critical band for both identity and expression shifted to slightly lower values as distance increased (Experiment 2), a pattern indicating only partial scale invariance. Children aged 10 and 14 years showed similar tuning but needed more contrast (Experiment 3). The patterns suggest that adults use finer details for recognizing facial expressions than for identifying faces, a tuning that appears as early as age 10.

1. Introduction

Facial identity and facial expression represent invariant and changeable aspects of faces, respectively. Human adults are fast and accurate in extracting these two types of information from faces. Several researchers have proposed that the recognition of facial identity and facial expression involve two separable systems (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000). That proposal is supported by evidence from behavioral measures (Young, McWeeny, Hay, & Ellis, 1986), neuropsychological studies (Etcoff, 1984; Hornak, Rolls, & Wade, 1996; Tranel, Damasio, & Damasio, 1988; Young, Newcombe, de Haan, Small, & Hay, 1993), functional imaging (George et al., 1993; Sergent, Ohta, Macdonald, & Zuck, 1994; Winston, Henson, Fine-Goulden, & Dolan, 2004), and single cell-recordings (Hasselmo, Rolls, & Baylis, 1989). However, no previous study has investigated whether we use the same or different spatial frequency information to recognize facial identity and facial expression.

Research on the recognition of facial identity has revealed that adults use a limited range of mid spatial frequencies, with spatial frequency defined as the number of sinusoidal transitions across the face, measured in cycles per face width (c/fw), rather than the number of variations across the retina, which is measured in cycles per degree. Unlike cycles per degree, cycles per face width remain constant as the face is viewed from different distances. Although it is generally agreed that the mid spatial frequency bands contain the critical information for face identification, the estimates of the critical center frequency vary from study to study, probably because of different ways of manipulating the available spatial frequency information.

Pixelizing was one of the earliest approaches used to manipulate spatial frequency information in faces. In this approach, a grid was put on a face image and the gray level within each block (pixel) was set to the mean gray level of the block. Studies using this approach reported that accuracy for recognition of facial identity dropped when the image quality dropped below 16 (Harmon, 1973), 18 (Bachmann, 1991), 21 (Costen, Parker, & Craw, 1994), and 23 (Costen, Parker, & Craw, 1996) pixels per face (8–11.5 c/fw). However, pixelizing introduces additional high spatial frequencies into the image, which may affect the estimates of the critical spatial frequency bands for face identification.

As a better way to manipulate spatial frequency information than pixelizing, filtering is a commonly used approach, because it allows the selective removal of certain spatial frequencies without adding extra spatial frequencies to the image. Using low-pass filtered faces, in which higher spatial frequencies were removed, Fiorentini, Maffei, and Sandini (1983) found that adults were less accurate in recognizing facial identity when the cutoff frequency dropped from 8 to 5 c/fw. Using low-pass and high-pass filtered faces, Costen and colleagues (1994, 1996) found that the most useful information for face identification is carried by a spatial frequency band between 8 and 16 c/fw. In contrast, using band-pass filtered faces, which only contain information in a narrow range

* Corresponding author. Address: Department of Psychology, Neuroscience & Behavior, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1. Fax: +1 905 529 6225.
E-mail address: maurer@mcmaster.ca (D. Maurer).

0042-6989/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2011.01.011

of spatial frequencies, Hayes, Morrone, and Burr (1986) found that the most critical information is located around 20 c/fw.

An alternative approach to selectively removing certain bands of spatial frequency by filtering is to add noise in the target spatial frequency band, a procedure called noise masking. Using narrow-band additive white Gaussian noise, Näsänen (1999) found that adults asked to recognize the identity of faces are most sensitive to spatial frequency information centered around 8–11 c/fw. Similar to noise masking, Fourier phase randomization selectively disrupts information in a certain spatial frequency band by scrambling the phase information in that band. This method has the advantage of leaving the amplitude spectrum of the face constant because no noise is added to the image. Using Fourier phase randomization, Näsänen (1999, Experiment 2) reported similar results (8–11 c/fw) to those found with noise masking (Näsänen, 1999, Experiment 1, 8–11 c/fw) for the critical spatial frequency band used by adults in face identification. Also using Fourier phase randomization, Ojanpää and Näsänen (2003) reported similar results (8–11 c/fw) when a visual search paradigm was used.

The use of spatial frequency information in face identification is assumed to be scale invariant. Hayes and colleagues (1986) found that the most informative spatial frequency band is located around 20 c/fw when adults were tested at 2.1 m or 8.5 m. However, Näsänen (1999, Experiment 4) and Ojanpää and Näsänen (2003) reported a slight shift of the critical spatial frequency band when adults were tested at different distances. When the testing distance was increased from 60 cm to 240 cm in the first study, the center of the critical spatial frequency band shifted from 11 c/fw to 6.9 c/fw. When the testing distance was increased from 57 cm to 171 cm in the second study, the center of the critical spatial frequency band shifted from 8–11 c/fw to 5.6–8 c/fw. One explanation of the shift is the attenuation of high spatial frequencies by the optics of the eye (Näsänen, 1999; Ojanpää & Näsänen, 2003). Nevertheless, when compared to the shift in retinal spatial frequency with increased viewing distance (e.g., from 60 cm to 240 cm in the Näsänen's (1999) study), the shift in object-based spatial frequency is relatively small. Therefore, the critical spatial frequency in recognizing facial identity appears to be partially scale invariant.

The far fewer studies on the recognition of facial expressions generally agree that the mid spatial frequency band is also critical. Using low-pass and high-pass noise masking, Schwartz, Bayer, and Pelli (1998) found that the critical spatial frequency band for recognizing facial expressions is located around 8 c/fw. Using hybrid faces, which contained a face with information below 8 c/fw superimposed on a face with information above 24 c/fw, Schyns and Oliva (1999) reported a low spatial frequency bias in categorizing facial expressions. The same low frequency bias was also found when the participants categorized facial identity in hybrid faces. Using filtered synthetic faces representing different facial expressions, Goren and Wilson (2006) found that when the spatial frequency band shifted from mid (10 c/fw) to low (3.3 c/fw) spatial frequency, discrimination thresholds increased for most of the expressions, especially for sadness, but not for anger. No change in threshold was found when the spatial frequency band shifted from mid (10 c/fw) to high (30 c/fw) spatial frequency. The difference in critical spatial frequency among expressions is also suggested by Smith and Schyns (2009)'s finding that different facial expressions are represented by different diagnostic spatial frequency spectra. In this study, the critical spatial frequency band for recognizing each expression revealed by the Bubbles technique (Gosselin & Schyns, 2001) was related, to some extent, to observers' sensitivity to that expression at different viewing distances. Expressions that have lower critical spatial frequencies (happy, surprise, disgust) were better recognized in smaller images simulating a longer distance than expressions that have higher critical spatial frequencies (neutral, sad).

Most previous studies reported only how human performance varied with experimental manipulations of the stimuli. In contrast, Näsänen (1999) compared the tuning curves of human observers and an ideal observer able to use all the pixel information contained in the images to distinguish facial identities. Unlike human observers, who show higher sensitivity to mid spatial frequencies than low or high spatial frequencies, for the ideal observer there was a linear decrease of sensitivity with increasing spatial frequency. The difference suggests that human observers are not able to use information in all spatial frequency bands equally well, that is, they are unable to use the greater information available in low spatial frequency bands as efficiently as the information in mid spatial frequency bands. Two previous studies (facial identity: Gold, Bennett, & Sekuler, 1999; facial expression: Schwartz et al., 1998) calculated ideal observer performance as a benchmark of information availability and used the ratio of ideal to human performance (efficiency) as a measure of the properties of human information processing. Both studies show that human observers are the most efficient in using the mid spatial frequency band in recognizing facial identity and facial expression.

There has not been a direct comparison of the critical spatial frequency bands for the recognition of facial identity and the recognition of facial expression in which the same methods were used to manipulate spatial frequency. Although previous studies agree that the mid spatial frequency band is critical for both facial identity and facial expression, the amount of variation across methods and studies makes it impossible to ascertain whether the critical bands are completely or only partially overlapping. That was the purpose of Experiment 1. We used noise masking to measure the critical spatial frequency band for the recognition of facial identity and of facial expressions, in each case with variable information from the other dimension. Specifically, we measured contrast thresholds for recognizing facial identity with varying expression and for recognizing facial expression with varying identity as a function of the spatial frequency of narrow-band additive white Gaussian noise. To increase the generality of the findings, we used four different identities and four facial expressions capturing the major variation in adults' perceptual structure of facial expressions (Gao, Maurer, & Nishimura, 2010), namely, happiness, sadness, fear, and anger. In Experiment 2, we used the same paradigm to investigate the effect of viewing distance. In Experiment 3, we used a subset of conditions to explore developmental changes in the critical spatial frequency band used in recognizing facial identity and expression.

2. General methods

2.1. Apparatus

The stimuli were generated on an Apple Mac Pro computer and displayed on a 21-in. CRT monitor (Dell P1130) with a resolution of 1600×1200, a refresh rate of 85 Hz (non-interlaced), and 256 grayscale levels. The average luminance of the stimuli and background was 20.4 cd/m^2. The experiments were controlled by custom software based on the Matlab (version 7.1) programming environment using Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Participants viewed the stimuli binocularly in a dimly lit room with their heads stabilized on a chinrest.

2.2. Face images

We used two female models (model number: 03, 10) and two male models (model number: 24, 25) from the Nimstim stimulus set (Tottenham et al., 2009). For each model, we selected the

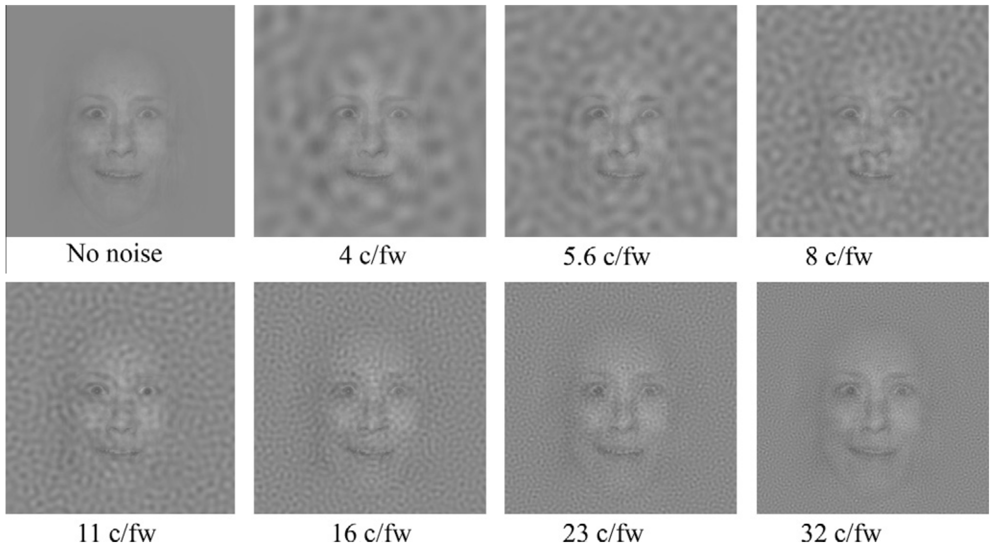

Fig. 1. Examples of faces without additive noise and with additive white Gaussian noise filtered at different center spatial frequency bands.

images depicting happy, sad, angry, fearful, and neutral expressions. The selected models received high agreement on the expression posed and high ratings of intensity from adults in a previous study (Palermo & Coltheart, 2004). The neutral faces and the original expressive faces were used only for generating the testing stimuli. For each expression, we created intensities of 50% and 90% by morphing the emotional face with the neutral face (for details, see Gao & Maurer, 2009), and only these two images were used during the test. The 90% expressions were near the maximum intensity produced by human adults and will be referred to as high intensity. The 50% expressions were lower but still quite a bit above threshold for each expression, as revealed in previous research with this stimulus set (Gao & Maurer, 2009). For ease, we will refer to them as low intensity. This procedure created 32 test stimuli (4 models × 4 expressions × 2 intensities).

Image processing was carried out using Matlab (version 7.1). The stimuli were converted to grayscale images and the amplitude spectrum of each face image was replaced by the average amplitude spectrum of the 32 face images. An oval-shaped Gaussian window was applied to each image to remove hair cues from the face (see Fig. 1). Each face had a width of 11 cm, or from a testing distance of 60 cm, 10.5 visual degrees.

2.3. Spatial frequency manipulation

On each trial, a white Gaussian noise mask that was the same size as the face image was superimposed. The noise mask was filtered by a Gaussian filter (see Supplementary Fig. 1 for the filter functions) at one of seven center frequencies (4, 5.6, 8, 11, 16, 23, 32 c/fw) with a bandwidth of 1.58 octaves (full width half height). The noise mask alone had a mean grayscale value of 0. When the noise mask was combined with the face image, the mean luminance value of the masked image was the same as the original face image. Image root-mean-square (RMS) contrast was computed by first computing local contrast:

$$c_i = \frac{l_i - L}{L} \tag{1}$$

where c_i is the contrast at pixel location i. L is average luminance and l_i is the luminance of the ith pixel. These values were then used to compute RMS contrast:

$$c_{\mathrm{RMS}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} c_i^2} \tag{2}$$

where n is the number of pixels in the image. The RMS contrast of the face image was varied according to a staircase procedure, while the RMS contrast of the noise was kept constant at 0.04 (the RMS contrast of the noise was reduced to 0.02 in Experiment 3). At a RMS contrast of 0.04, the standard deviation of the grayscale values of the noise mask was 5.47. Since there were only 256 gray levels, any gray scale values in the masked image larger than 255 or smaller than 0 were truncated to be within the range of 0 and 255.[1]

2.4. Ideal observer analysis

We built an ideal observer that had pixel information for all the original images (the templates, T) and information on the noise filter functions. On each trial, the stimulus image I (signal plus filtered noise) was Fourier transformed. The amplitude spectrum of I was divided by the magnitude spectrum of the filter function, which had been used as the noise in the current trail, while the phase spectrum of I was preserved. The new amplitude spectrum and the preserved phase spectrum were then submitted to inverse Fourier transform to generate $I_{C^{-1}}$. The resulting image $I_{C^{-1}}$ is equivalent to an inverse-filtered signal plus the original white Gaussian noise. This inverse-filtering process is called "prewhitening", which removes the correlations between noise amplitude spectrum and spatial frequency (Conrey & Gold, 2009). The ideal observer also inverse-filtered the templates ($T_{C^{-1}}$) using the same filter function to preserve the probability structure of the original image I relative to the ideal templates T. Then the ideal observer submitted $I_{C^{-1}}$ and $T_{C^{-1}}$ to a decision rule, in which the ideal observer calculated the ratio between the posterior probabilities that $I_{C^{-1}}$ was from stimulus group A or from stimulus group B using the following formula (Gold et al., 1999; Tjan, Braje, Legge, & Kersten, 1995):

[1] Truncating the noise can affect the normality of the noise distribution by introducing "spikes" to the ends of the distribution. However, this would not have any substantial effect on the current stimuli since at near threshold contrast level (RMS contrast of 0.1) the probability of having the grayscale value of one pixel falling out of the display range (0–255) in the entire image (650 × 650 pixels) is small ($p = 0.0003$).

$$p = \frac{\sum_{j=1}^{Na} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(I_{C_i^{-1}} - T_{C^{-1}}a_{ij})^2\right]}{\sum_{j=1}^{Nb} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(I_{C_i^{-1}} - T_{C^{-1}}b_{ij}\right)^2\right]} \quad (3)$$

where $I_{C_i^{-1}}$ is the grayscale value of the ith pixel of the prewhitened stimulus, $T_{C^{-1}}a_{ij}$ is the value at pixel i in the jth prewhitened template of stimulus group A, Na is the number of templates in stimulus group A, $T_{C^{-1}}b_{ij}$ is the value at pixel i in the jth prewhitened template of stimulus group B, Nb is the number of templates in stimulus group B, n is the number of pixels in each image (650 × 650), and σ is the standard deviation of the noise grayscale values.

On a given trial, the ideal observer chose answer A if p was greater than 1, and answer B if p was less than 1. The ideal observer ran all the conditions (except the no noise condition) with the same settings as human observers. For each condition the threshold was the mean threshold of 10 runs. We calculated the efficiency of human information processing in each condition as the ratio between the ideal observer threshold contrast energy (E_{ideal}) and the mean of human threshold contrast energy (E_{human}), and converted efficiency to relative sensitivity by calculating the logarithm of the reciprocal of efficiency.

$$\text{Efficiency} = E_{\text{ideal}}/E_{\text{human}} \quad (4)$$

Contrast energy is defined as:

$$E = c_{RMS}^2 na \quad (5)$$

where n is the number of pixels in the picture and a is the area of a single pixel.

2.5. Procedure

The protocol was approved by the McMaster Research Ethics Board. We obtained written consent from the adult participants or from a parent of the child participants, and we obtained verbal assent from the child participants.

2.5.1. Training

Participants began with training to recognize the identity of all the target faces that would be used in the following testing session. Each training task had three stages. In the first stage, participants passively viewed the eight versions of each target face identity (with four expressions × two intensities) twice, for 2 s each time, preceded by a label identifying the face (female 1, female 2, male 1, male 2). In the second stage, participants indicated the identity of each face image by pressing a predefined key. Each version of the target face images was shown once for 500 ms (the presentation time was extended to 1000 ms in Experiment 3). Auditory feedback was given after each key press with a high pitch tone indicating a correct response and a low pitch tone indicating an incorrect response. The third stage was identical to the second stage except that no feedback was given after each response. Participants needed to reach 100% accuracy in the third stage to proceed to the testing session. No training was given for the facial expression discrimination task.

2.5.2. Testing

Each trial started with a 500 ms presentation of a fixation cross in the center of the screen followed by a 500 ms presentation of a face image (the presentation time was extended to 1000 ms in Experiment 3). Participants used the keyboard to indicate their answers and received the same auditory feedback as used in the training session. The next trial began as soon as the feedback ended. On the first trial of each staircase, the face image had a RMS contrast of 0.2 (a RMS contrast of 0.3 was used for the first trials in Experiment 3). After three (two for Experiment 2 and Experiment 3) consecutive correct responses, the RMS contrast of the face image was decreased by a factor of 1.26. After each incorrect response, the RMS contrast of the face image was increased by the same factor. The staircase procedure terminated at 80 trials or 10 reversals, whichever came first. The threshold value was calculated as the geometric mean of the RMS contrast values of the last six reversals, representing an accuracy of 0.79 (0.71 for Experiment 2 and Experiment 3).

3. Experiment 1

In Experiment 1, we examined the importance of different bands of spatial frequency for recognizing facial identity and facial expression by adding to each face image, white Gaussian noise that masked one of seven narrow spatial frequency bands with different center frequencies. In the facial identity task, the observers learned and discriminated a pair of male faces and a pair of female faces, in each case with varying facial expressions. In the facial expression task, the same observers discriminated two pairs of facial expressions (happiness vs. sadness and anger vs. fear) posed by the same four models. Each expression pair was tested at both high intensity (90%) and low intensity (50%). We chose these two pairings of facial expression because they differ maximally on the two major dimensions in adults' perceptual structure for facial expressions, namely, the pleasure dimension and the potency dimension (Gao et al., 2010). Happiness and sadness represent two ends of the pleasure dimension, while anger and fear represent two ends of the potency dimension. We also compared the performance of the five observers to the performance of an ideal observer with all the pixel information in the images.

3.1. Participants

Participants were five adult observers (XG, SM, JW, MV, and OK, age range: 20–28) from McMaster University. XG and MV are experienced psychophysical observers. SM, JW, and OK were naïve to the purpose of the current study and had very limited previous experience in psychophysical experiments. All participants had normal or corrected-to-normal vision. They completed testing over 2 weeks.

3.2. Design

3.2.1. Facial identity discrimination

Participants discriminated between the two male models in one block and between the two female models in the other block; order of blocks was randomized across observers. For each model there were eight different images consisting of four expressions at two intensities. In each block, two staircase runs were conducted for each of the eight noise masking conditions (no noise and seven center frequencies). In total, each participant generated 32 thresholds (two testing blocks [male; female] × eight noise conditions × two runs). The order of testing within each block was constrained so that the first and the ninth staircase were always the no noise condition and the other conditions appeared in a random order once before and once after the ninth staircase. The mean threshold of the two runs was used as the threshold for each condition.

3.2.2. Facial expression discrimination

The facial expression task consisted of four blocks: 90% happiness vs. 90% sadness, 50% happiness vs. 50% sadness, 90% anger vs. 90% fear, and 50% anger vs. 50% fear. Since we used the same four models for both identity and expression discrimination, we always ran the identity task before the expression task to control the amount of exposure to the four identities. Within each block, the order of masking conditions was controlled in the same way as for the identity discrimination task. In total, each participant generated 64 thresholds (four testing blocks × eight noise conditions × two runs). The four blocks of expression discrimination were tested in a pseudo random order so that the low intense pairings always followed their corresponding high intense pairings, with the starting expression pair (happy/sad or angry/fear) randomized across observers. The mean threshold for the two runs was used as the threshold for each condition.

3.3. Results and discussion

3.3.1. Contrast thresholds

Fig. 2a indicates that the peaks of the tuning functions as a function of the masked spatial frequency band for the two facial identity discrimination tasks (the solid lines), one involving two male faces and the other, two female faces, are at 11 c/fw. Since we used a noise masking paradigm, higher contrast threshold suggests that the masked information is more important for the task. Therefore, the figure suggests that human observers make the most use of information in the spatial frequency band around 11 c/fw when discriminating facial identity. This value is consistent with previous findings that human adults rely on mid spatial frequencies (8–16 c/fw) to recognize facial identity (Costen et al. (1994, 1996), Fiorentini et al. (1983), Gold et al. (1999), Näsänen (1999), Ojanpää and Näsänen (2003); for review, see Ruiz-Soler & Beltran, 2006). The patterns for facial expression are different from those for facial identity, although they were consistent across the two discriminations (happiness vs. sadness and anger vs. fear) and two intensities (90% and 50%). In most cases, the peak of the tuning functions (the dashed lines) in the facial expression discrimination tasks is at 16 c/fw.

To quantity the masking functions, we fit Gaussian functions to the group mean for each task using the following formula:

$$y = A \times \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \quad (6)$$

Results and discussion

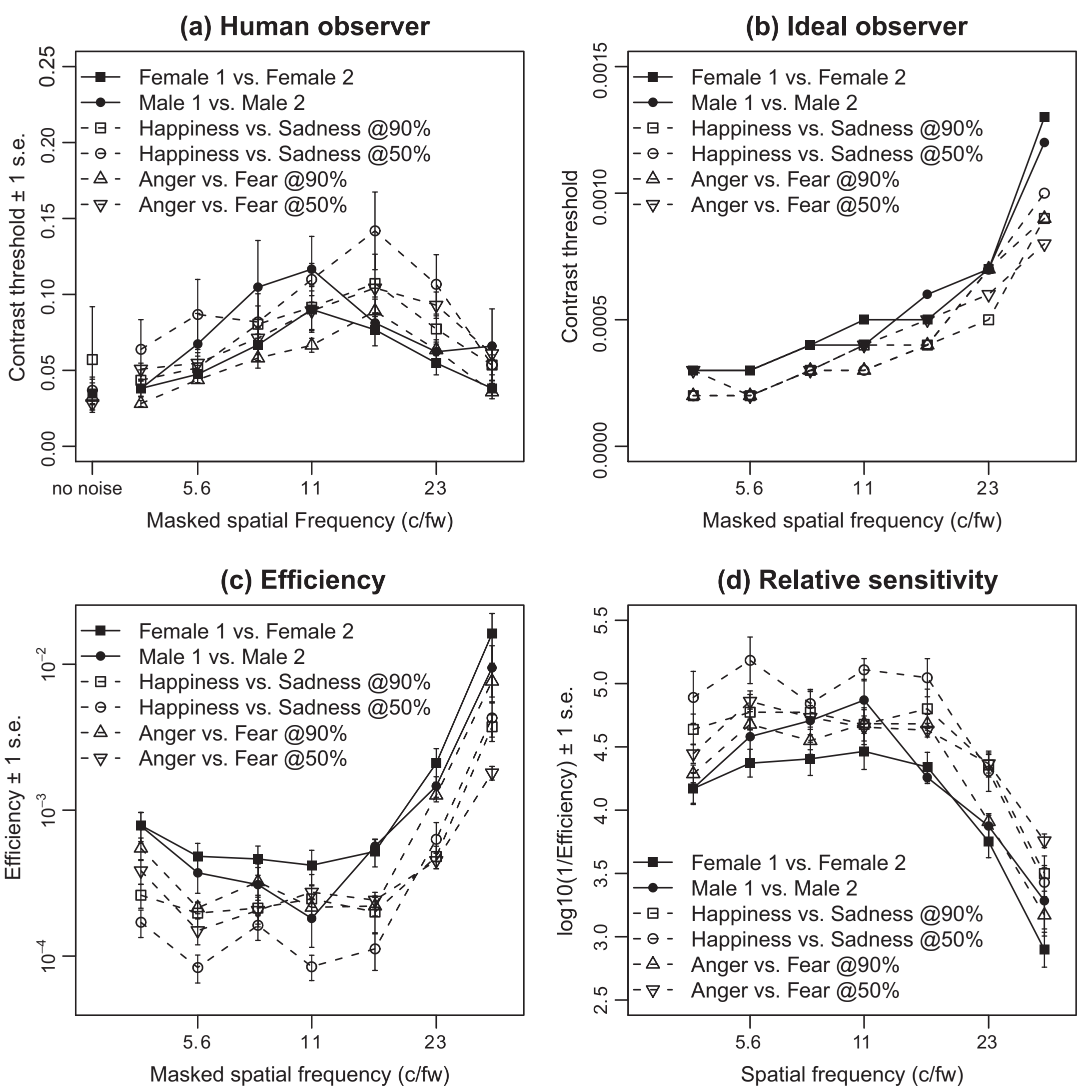

Fig. 2. (a) RMS contrast threshold as a function of the center spatial frequency of the masking noise for human observers for each of the six tasks. Higher values indicate that performance dropped when that spatial frequency band was masked and hence that human observers rely more on that band for the task. (b) RMS contrast thresholds for the ideal observer plotted in the same way. Higher values indicate that the performance of the ideal observer dropped when that spatial frequency band was masked. (c) Efficiency. Ratio of ideal to human observer threshold contrast energy for each center frequency of masking noise. (d) Relative sensitivity as a function of spatial frequency. Plotted is the inverse of efficiency compared to the ideal observer so that higher values indicate greater sensitivity.

where x is spatial frequency in log units, y is the contrast threshold value, A is the height of the peak, μ is the position of the peak, and σ represents the width of the curve. We estimated the position of the peak (μ) for each task using a bootstrapping procedure based on 1000 iterations. The estimated values indicate that the peak masking is around 11 c/fw for facial identity, a value significantly lower ($ps < .05$) than the peak, around 13 c/fw, for facial expression (see Supplementary Table 1 for the estimates of peak positions).

3.3.2. Ideal observer

Before we attribute the different spatial frequency tuning to the difference in the ability of human observers to use information in processing facial identity and facial expressions, we have to consider the availability of information in these tasks. As shown in Fig. 2b, the ideal observer performed similarly on the two types of task, in each case much better than human observers as indicated by lower contrast thresholds. Unlike human observers, whose tuning curves have an obvious peak when the masked spatial frequency is in the mid spatial frequency range, the tuning curves of the ideal observer show an increase in contrast thresholds as the masked center spatial frequency increases. To test whether the shapes of the tuning curves of the ideal observer are the same for different tasks, we first found the best fitting function to the means across all six tasks. A quadratic function $f(x) = p_1x^2 + p_2x + p_3$ ($p_1 = 0.0012$, $p_2 = -0.0018$, $p_3 = 0.0009$) provided a good fit ($r^2 = 0.98$). We then fit a quadratic function to the mean of each task using the same p_1 and p_2 values as estimated in the previous step but allowing p_3 to vary since p_3 only affects the height but not the shape of the curves. A one-way ANOVA on the residuals of the fits to the six tasks revealed no significant difference, $F(5, 59) = 1.587$, $p = 0.179$, suggesting that the shapes of the tuning functions for different tasks are not different from each other. A comparison of Fig. 2a and b suggests, like previous studies (Gold et al., 1999; Näsänen, 1999), that human observers do not only rely on low-level image information to discriminate either facial identity or facial expression: their maximum sensitivity is at a different masked spatial frequency and, unlike the ideal observer, differs for expression and identity judgments.

3.3.3. Relative sensitivity

Using the performance of the ideal observer as a benchmark, we gauged human performance by calculating human efficiency (Fig. 2c) using formula 4. Since we used a masking paradigm in the current study, the lowest efficiency value indicates the spatial frequency band that is most important because it leads to the largest drop in efficiency when it is masked. To present human sensitivity in a more intuitive way (higher value represents higher sensitivity), we calculated and plotted relative sensitivity as the logarithm of the reciprocal of the efficiency (Fig. 2d). For all six tasks, human sensitivity is high in a wide range running from low (4 c/fw) to mid (11–16 c/fw) spatial frequencies and drops rapidly beyond the mid (11–16 c/fw) spatial frequency band. Unlike the curves of contrast thresholds (Fig. 2a), there is no obvious peak in the curves of relative sensitivity. Although the shapes of the curves of relative sensitivity are similar across tasks, the positions of the "elbows" of the curves (the spatial frequency where sensitivity begins to drop) are at a slightly lower spatial frequency for the identity than the expression tasks (Fig. 2d, solid versus dashed lines). To quantify this difference, we fit a step function to the mean across the two identity tasks and the mean across the four expression tasks:

$$f(x) = \begin{cases} c & \text{if} \quad x \leqslant (c-b)/a \\ ax+b & \text{if} \quad x > (c-b)/a \end{cases} \tag{7}$$

where $(c - b)/a$ is the position of the "elbow". The "elbow" of the curve for the mean of the two identity tasks is at 17.4 c/fw (95% bootstrapped confidence interval: 15.1–18.4 c/fw), which is lower than for the mean of the four expression tasks (18.5 c/fw, 95% bootstrapped confidence interval: 18–18.8 c/fw). Therefore, the lower critical spatial frequencies for the facial identity than for the facial expression task revealed by the contrast thresholds can be explained by the earlier drop of sensitivity at high spatial frequencies when judging identity than when judging expression.

The results for Experiment 1 indicate that adult human observers are most affected by masking of mid spatial frequencies when judging facial identity or expression, but that when the information available in the stimuli, as revealed by ideal observer analysis, is taken into account their sensitivity is equally high for spatial frequencies running from low to mid band levels. Whether based on the original contrast thresholds or their sensitivity relative to an ideal observer, they are less well able to use information at higher spatial frequencies, with an increasing drop-off as spatial frequency increases that begins earlier for facial identity than for facial expression. Thus, as the combined result of sensitivity and information availability, the most useful spatial frequencies for adults discriminating happiness vs. sadness and fear vs. anger includes a spatial frequency band that is higher than the most useful bands for recognizing facial identity, whether the facial expression is high or low in intensity. The current finding adds a new piece of evidence to the proposal (Bruce & Young, 1986; Haxby et al., 2000) that in adults, the systems for processing facial identity and facial expression are at least partially separate.

4. Experiment 2

Previous studies of the recognition of facial identity have reported that the critical spatial frequency band is largely constant across distance, that is, it is based on the amount of information in each unit of the face (cycles per face width), rather than in each unit of the retina (cycles per degree) (Hayes et al., 1986; Näsänen, 1999; Ojanpää & Näsänen, 2003). Such an object-based system is useful for recognizing faces across varying distance. However, the empirical results for facial identity have been inconsistent: when distance was varied 3–4-fold, one study found no change in the optimal spatial frequency over a fourfold increase (Hayes et al., 1986), while two others found a shift toward a slightly lower spatial frequency band when distance was increased (Näsänen, 1999; Ojanpää & Näsänen, 2003). There has been no report on the effect of distance on the critical spatial frequency band for the recognition of facial expressions.

In Experiment 2, we investigated the critical spatial frequency band for the recognition of facial identity and of facial expressions at three testing distances using the same noise masking paradigm as in Experiment 1. Because in Experiment 1, the tuning functions were similar for discriminating the two females and the two males and similar for the four facial expression tasks, we used two four-alternative forced choice tasks in Experiment 2: one with the four identities and one with the four expressions at the high intensity (90%).

4.1. Participants

Participants were the same five observers as in Experiment 1. They began Experiment 2 after completing Experiment 1 and completed it over a number of days within a 2-week period.

4.2. Design and procedure

4.2.1. Facial identity discrimination

In the identity block, participants discriminated among the four models (two female, two male) used in Experiment 1. For each model, there were eight different images consisting of four expressions at two intensities (50% and 90%). Participants were tested in sequence at three viewing distances: 60 cm, 120 cm, and 180 cm. For each testing distance, two staircase runs were conducted at each of the eight noise masking conditions to yield 48 thresholds (three viewing distances × eight noise conditions × two runs). The order of the conditions was controlled in the same way as in Experiment 1. The mean threshold of the two runs was used as the threshold for each condition.

4.2.2. Facial expression discrimination

In the expression block, participants discriminated among four expressions: happiness, sadness, fear, and anger, all at 90% intensity. Each expression was displayed in the faces of the same four models used in Experiment 1. The other details of the procedure were the same as in the identity discrimination task. We collected 48 thresholds (three viewing distances × eight noise conditions × two runs) from each participant. The expression block was always run after the identity block.

4.3. Results and discussion

4.3.1. Contrast thresholds

As shown in Fig. 3a, the contrast threshold functions have obvious peaks when the mid spatial frequency range is masked. To quantify the positions of these peaks, we ran the same bootstrapping procedure as in Experiment 1 (see Supplementary Table 2 for the estimates of peak positions). The estimates indicate that at a viewing distance of 60 cm, the peak sensitivity for facial expression discrimination is at a higher masked spatial frequency (12.5 c/fw, 95% bootstrapped confidence interval: 11.9–12.9 c/fw) than for facial identity discrimination (11.4 c/fw, 95% bootstrapped confidence interval: 10.9–12.2 c/fw). This pattern is also evident at the two other testing distances. Thus, regardless of distance, adult observers use slightly higher spatial frequency information (i.e., finer details) to distinguish facial expressions than they do to differentiate identity. The estimates also indicate that when the viewing distance increased to 120 cm and 180 cm, the peak sensitivity for both facial identity discrimination and facial expression discrimination moved to lower masked spatial frequencies (from 11.4 to

Results and discussion

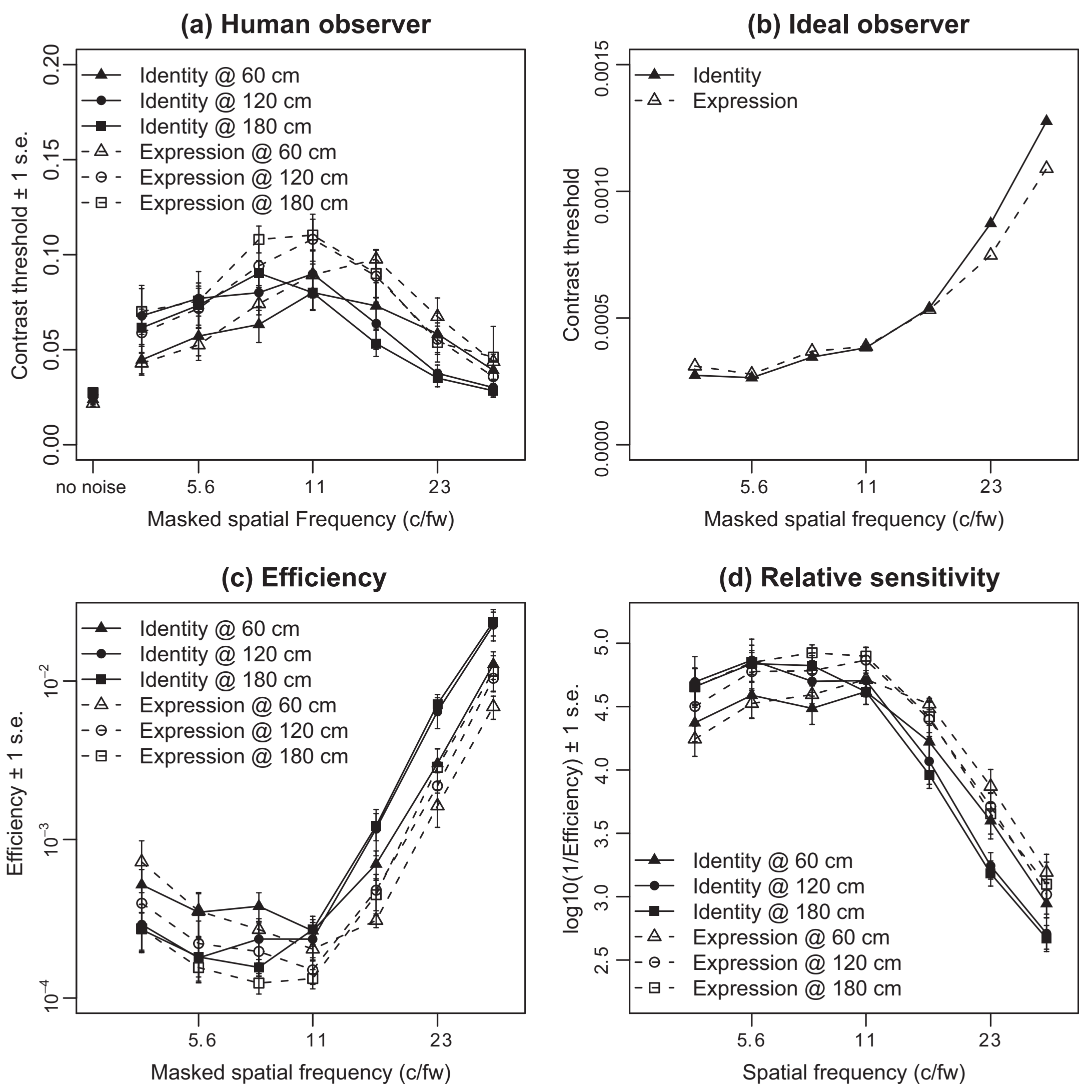

Fig. 3. (a) RMS contrast threshold as a function of the masked center spatial frequency of the masking noise for human observers for each task and distance. Higher values indicate that performance dropped when that spatial frequency band was masked and hence that human observers rely more on that band for the task. (b) Performance of the ideal observer for the two tasks as a function of masked spatial frequency. Higher values indicate that the performance of the ideal observer dropped when that spatial frequency band was masked. (c) Efficiency. Ratio of ideal to human observer threshold contrast energy for each center frequency of masking noise. (d) Relative sensitivity as a function of spatial frequency. Plotted is the inverse of efficiency compared to the ideal observer so that higher values indicate greater sensitivity.

around 8 c/fw for identity and from 12.5 to around 10 c/fw for expression).

4.3.2. Ideal observer

The ideal observer behaved similarly for the identity task and the expression task in Experiment 2 (Fig. 3b). A quadratic function $f(x) = p_1x^2 + p_2x + p_3$ with the same values of p_1 and p_2 as used in Experiment 1 provided a good fit to the data for both the identity task and the expression task ($r^2 = .91, .98$, respectively). Therefore, despite the difference in task format (four-alternative forced choice in Experiment 2 vs. two alternative forced choice in Experiment 1), the ideal observer showed almost identical patterns across the two experiments.

4.3.3. Relative sensitivity

Fig. 3d shows relative sensitivity calculated in the same way as in Experiment 1. The patterns are very similar to those in Experiment 1. For all the conditions, relative sensitivity is high in the low to mid spatial frequency range and drops beyond mid spatial frequencies. We estimated the positions of the "elbows" in the same way as in Experiment 1. As shown in Table 1, at the same testing distance, the position of the "elbow" for the identity task is at a lower spatial frequency than for the expression task. For both tasks, when distance increased the "elbows" moved to a lower spatial frequency.

The shift of the "elbow" position of the relative sensitivity functions is smaller than would be expected if it were based on retinal image size: the position of the elbow for facial identity in retinal coordinates at 60 cm was at 1.3 cycle/degree, which corresponds to 4.3 c/fw at 180 cm; for facial expression, the corresponding figures are 1.4 cycle/degree at 60 cm and 5.3 c/fw at 180 cm. The actual shifts observed are much smaller than predicted based on the retinal image and suggest that instead, the critical spatial frequency information used for both facial identity discrimination and for facial expression discrimination is largely object-based.

Unlike the current results, Hayes et al. (1986) found a higher than typical optimal spatial frequency band (20 c/fw) for distinguishing four facial identities at both 2.1 m and 8.5 m. The higher value may have arisen because, unlike other studies, the faces included hair cues that might be more easily distinguished with information from a higher band of spatial frequencies than is optimal for discriminating internal features and face shape and/or because the faces were presented at a much lower luminance (8 cd/m^2 versus 20.4 cd/m^2 in the current study). Also important might be differences between studies in stimulus size and distances tested. The retinal face sizes (16 visual degrees at 2.1 m, 4 visual degrees at 8.5 m) in Hayes et al. (1986) were slightly larger than in other studies (10.5 visual degrees at 60 cm, 3.8 visual degrees at 1.8 m [current study]; 9.5 visual degrees at 60 cm, 2.4 visual degrees at 2.4 m [Näsänen, 1999]; 3.2 visual degrees at 57 cm, 1.1 visual degrees at 1.7 m [Ojanpää & Näsänen, 2003]). The "near" distance (2.1 m) in Hayes et al. (1986) was equivalent to the farthest distance in other studies (1.8 m [current study], 2.4 m [Näsänen, 1999], 1.7 m [Ojanpää & Näsänen, 2003]) and the drop in optimal spatial frequency with viewing distance may decrease with variations beyond 2 m. In fact, as indicated in Table 2, most of the drop-off in the current study was between .6 and 1.2 m, with little further decrease at 1.8 m. Thus, beyond 1–2 m, the critical spatial frequency band for facial identity and facial expression may be completely object-based and perfectly scale invariant. However, our data indicate that for faces at distances less than 1 meter, the critical band moves to (slightly) higher spatial frequencies as the face moves closer to the observer for judgments of both identity and expression.

Such object-based scaling of the critical spatial frequency information is useful in real life because faces are seen at different distances, resulting in different retinal image sizes, but the critical information distinguishing them and their facial expressions is largely constant at an object-based scale. The small deviation from perfect scale invariance is consistent with findings from previous studies (Näsänen, 1999; Ojanpää & Näsänen, 2003) that with increasing viewing distance, the peak sensitivity for facial identity moves to a slightly lower object-based spatial frequency and extends those findings to the discrimination of facial expression. The optical attenuation of high spatial frequencies at greater distances is a possible explanation for the change (Näsänen, 1999; Ojanpää & Näsänen, 2003).

Besides the optical attenuation of high spatial frequencies at greater distances, the shape of the contrast sensitivity function may also contribute to the drop in the optimal spatial frequency with increasing distance—at greater distances, the highest spatial frequency that is visible to the subject will shift to lower object-based values. However, the findings do not fit this prediction. For facial identity, the elbow, the spatial frequency where sensitivity begins to drop-off is around 14 c/fw at 60 cm (4 cycles/degree at 180 cm). At 180 cm, the elbow is around 10 c/fw, which corresponds to 2.9 cycles/degree, a shift (4–2.9 cycles/degree) moving away from the peak of the contrast sensitivity function. The same pattern is observed for facial expressions: the elbow shifted from 16 c/fw at 60 cm (4.4 cycles/degree at 180 cm) to 13 c/fw at 180 cm (3.7 cycles/degree at 180 cm).

5. Experiment 3

Experiments 1 and 2 showed that adults use information mainly from low and mid spatial frequencies to recognize facial identity and facial expression and that the critical band is lower for identity than expression. Little is known about how children use spatial frequency information in face perception. Contrast sensitivity is adult-like by age 7 (Ellemberg, Lewis, Maurer, Liu, & Brent, 1999; but see Benedek, Benedek, Kéri, & Janáky, 2003, for continued change until age 11–12). However, at age 10, children

Table 1
Bootstrapped estimates of the "elbow" position.

Task/distance (cm)	"elbow" position (c/fw)	95% confidence interval of "elbow" position	
		Lower band	Upper band
Identity			
60	13.7	12.9	14.9
120	10.8	8.6	12.7
180	10.1	7.4	11.5
Expression			
60	16.8	15.5	18.7
120	13.6	10.4	14.8
180	12.7	12.0	13.2

Table 2
Bootstrapped estimates of the "elbow" position.

Age/task	"elbow" position (c/fw)	95% confidence interval of "elbow" position	
		Lower band	Upper band
10 years			
Identity	11.8	10.3	14.2
Expression	14.2	12.0	13.2
14 years			
Identity	12.5	9.5	14.5
Expression	14.0	13.2	14.9
Adults			
Identity	11.7	9.6	12.9
Expression	14.1	13.4	14.6

still have higher thresholds than adults for recognizing some of the basic emotional expressions (Gao & Maurer, 2009) and even at age 14, they judge the similarity of facial expressions less categorically than adults (Gao et al., 2010). Children's accuracy in discriminating facial identity continues to improve after age 10–14, at least for faces differing only subtly in the spacing of features (Mondloch, Dobson, Parsons, & Maurer, 2004; Mondloch, Geldart, Maurer, & Le Grand, 2003; Mondloch, Le Grand, & Maurer, 2002). Some of this late improvement might be related to using a different, and less optimal, band of spatial frequencies for discrimination than used by adults, or to lower sensitivity within the critical band. Only two previous studies have investigated spatial frequency tuning in face perception in typically developing children. Deruelle and Fagot (2005) compared children aged 5–8 years to adults using a hybrid face paradigm similar to Schyns and Oliva (1999) in which a face filtered to have spatial frequencies only below 12 c/fw was superimposed on a face filtered to have frequencies only above 36 c/fw. The two superimposed faces with different filtering differed in identity or in expression (smile or grimace). Adults judged facial identity from the low-pass face and showed no bias when asked to judge whether the facial expression was a smile or a grimace, perhaps because neither choice matched their critical frequency band (cf. Experiment 1). Children also judged facial identity from the low-pass face but, unlike adults, judged facial expression from the high-pass face. It is difficult to interpret these results because only two fairly large spatial frequency bands were contrasted, because the informative mid spatial frequency band was not included, and because the faces were presented for a sufficiently long duration (400 ms for children and 100 ms for adults compared to 50 ms in the original study by Schyns and Oliva (1999)) that the participants may have been able to process both faces in the hybrid image and use more analytic higher-level strategies to make the decision. Using a masking paradigm, Leonard, Karmiloff-Smith, and Johnson (2010) found that, like adults, but unlike younger children, 9- and 10-year-olds are more sensitive to mid spatial frequency bands (centered on 16 c/fw) than to either low (8 c/fw) or high (32 c/fw) spatial frequency bands in recognizing the identity of upright faces, with no such bias for inverted faces. Facial expressions were not included in the study, thresholds were not measured, and no ideal observer analysis was included in order to calculate efficiency.

In Experiment 3, we used the noise masking paradigm to compare children at age 10 and 14 years to adults on the critical spatial frequency band for facial identity and facial expression. Specifically, we adapted the methods used in Experiments 1 and 2 to be suitable for children by reducing the number of conditions, reducing the contrast of the noise, and presenting the faces for a longer duration (see general methods for the details of the changes). Instead of using a broad range of spatial frequency bands as in Experiments 1 and 2, we used only four spatial frequencies bracketing the critical band for adults, namely, 5.6, 11, 16, and 32 c/fw. Each participant completed a block of judgments of facial identity involving either the two females or the two males and a block of judgments of facial expressions involving the highly intense (90%) happy and sad expressions.

5.1. Participants

The final sample included 16 10.5-year-olds (±3 months), 16 14-year-olds (±3 months), and 16 adults who had not participated in Experiments 1 and 2 (18–28 years of age, mean = 19.6). Child participants were recruited from names on file of parents who had volunteered their children at birth for participation in later studies. Adults were undergraduate students participating for course credit. Half of the participants in each age group were female. All of the participants had normal or corrected-to-normal vision. An additional five participants were excluded from data analysis because they failed visual screening (two 10-year-olds) or failed to reach the criterion in the training session (two 14-year-olds and one adult).

5.2. Design

5.2.1. Facial identity discrimination

Half of the participants in each age group of each sex discriminated between the two male models, while the other half discriminated between the two female models. For each model there were 8 different images consisting of four expressions at two intensities. One staircase was run at each of the four noise masking conditions (5.6, 11, 16, 32 c/fw) to collect four thresholds for each participant.

5.2.2. Facial expression discrimination

Participants discriminated between 90% happiness and 90% sadness. As in Experiments 1 and 2, each expression was displayed on the faces of four different models. One staircase was run at each of the four noise masking conditions to collect four thresholds for each participant.

Half of the participants in each age group of each sex were tested on the facial identity task first, while the other half of participants were tested on the facial expression task first. The order of the spatial frequency conditions within each task was controlled by a Latin square design. Participants were tested at a viewing distance of 60 cm.

5.3. Results and discussion

5.3.1. Contrast thresholds

Fig. 4a shows the mean contrast thresholds in each condition. Consistent with Experiments 1 and 2, in adults, the peaks of the masking curves for recognizing identity and expression are located at 11 and 16 c/fw, respectively. The estimated peak positions (see Supplementary Table 3 for the bootstrapped estimates of peak positions) also replicate Experiments 1 and 2 in indicating that adults' peak sensitivity for recognizing facial expression is at a higher spatial frequency than for recognizing facial identity.

As shown in Fig. 4a, contrast thresholds to recognize facial identity and facial expressions decrease with age, $F(2, 45) = 15.2$, $p < .01$. There is a significant reduction in threshold between 10 and 14 years of age ($p < .01$) and another significant reduction between 14 years of age and young adulthood ($p < .05$). In the 10- and 14-year-olds, the peaks of the masking curves for recognizing identity are located at 11 c/fw, as they are in adults. However, in the 10- and 14-year-olds, there is no obvious peak for the masking curves of recognizing expression. Nevertheless, the bootstrapped estimates indicated similar peaks for each task across the three age groups, with the only difference being that, for both the 10- and 14-year-olds, the estimated peaks for identity and expression are not statistically different from each other.

5.3.2. Ideal observer

The ideal observer shows similar pattern for the two tasks, as in Experiments 1 and 2. The same quadratic function used in Experiments 1 and 2 fit well the curves for the identity task and the expression task ($r^2 = .91, .93$, respectively).

5.3.3. Relative sensitivity

From Fig. 4d, we can see the similar pattern for all three age groups for both tasks, which is, in turn, similar to that seen in Experiments 1 and 2. The relative sensitivity is high in the low to mid spatial frequency range for both tasks at all ages and drops beyond mid spatial frequencies. Using the same procedure as in Experiments 1 and 2, we estimated the position of the "elbows"

Results and discussion

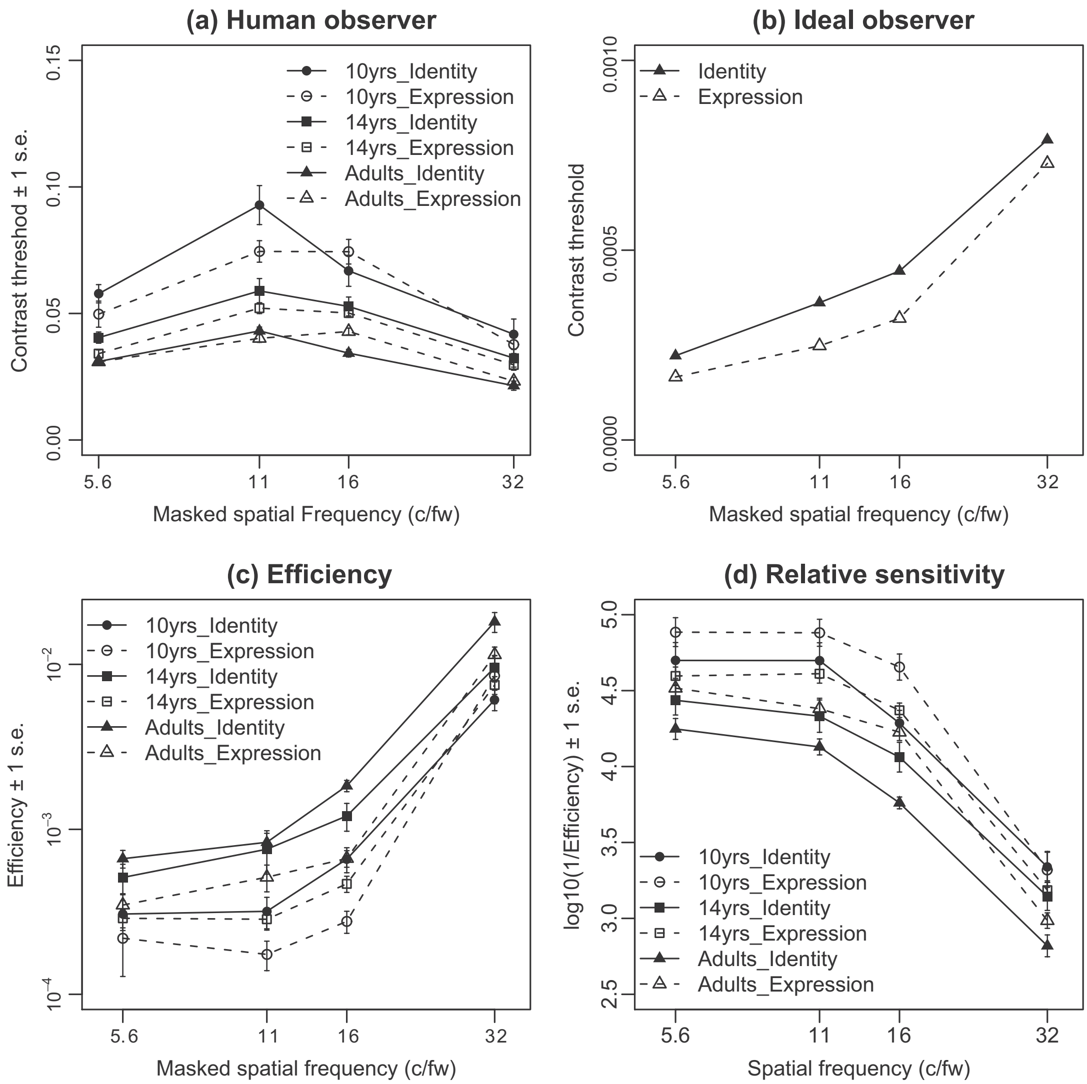

Fig. 4. (a) RMS contrast threshold as a function of masking noise spatial frequency for each age group on each task. Higher values indicate that performance dropped when that spatial frequency band was masked and hence that human observers rely more on that band for the task. (b) RMS contrast threshold as a function of masking noise spatial frequency for the ideal observer. Higher values indicate that the performance of the ideal observer dropped when that spatial frequency band was masked. (c) Efficiency. Ratio of ideal to human observer threshold contrast energy for each center frequency of masking noise. (d) Relative sensitivity as a function of spatial frequency. Plotted is the inverse of efficiency compared to the ideal observer so that higher values indicate greater sensitivity.

of the curves (Table 2). Consistent with Experiments 1 and 2, adults' relative sensitivity to facial identity drops at a lower spatial frequency than their sensitivity to facial expression. Children at age 10 and 14 have lower sensitivity overall but the same pattern as adults.

The current findings suggest that children as young as 10 years of age use the available spatial frequency information in the same way as adults in recognizing facial identity and facial expressions. However, 10-year-olds need a great deal more contrast than adults to recognize facial identity and facial expressions, a difference that diminishes but does not disappear by age 14. This developmental difference may apply to the recognition of object characteristics in general and arise from differences in ability to extract visual signals from noise or differences in general attentional or high-level visual abilities (e.g., contour integration continues to improve past age 14, Kovács, Kozma, Fehér, & Benedek, 1999: Hadad, Maurer, & Lewis, 2010). The consequence is that children as old as 14 years of age will have more trouble than adults in recognizing faces in poor light.

The current findings for facial identity are consistent with previous reports that 5- and 8-year-olds (Deruelle & Fagot, 2005) and 9–10 year-olds (Leonard et al., 2010), like adults, do not rely on high spatial frequency information (>36 c/fw or 32 c/fw, respectively) to recognize facial identity. They also agree with one of those studies (Leonard et al., 2010) that 10-year-olds, like adults, use mid spatial frequency information relatively efficiently when judging identity. However, by using ideal observer analysis and calculating relative sensitivity, our results show that, in addition, 10-year-olds, like adults, are able to use low spatial frequently information efficiently in judging facial identity.

The results for facial expression contrast with the one previous developmental study. Deruelle and Fagot (2005) found that 5- and 8-year-olds rely on high spatial frequencies (>32 c/fw) to recognize facial expression, while we found that 10- and 14-year-olds, like adults, are most efficient at using information below 14 c/fw to recognize facial expression. The difference in results may simply reflect developmental changes between 8 and 10 years. However, it is possible that the difference arises from the different methods used in these two studies. The previous study (Deruelle & Fagot, 2005) left out a large part of the mid spatial frequency range (12–20 c/fw), which previous studies, like this one, have shown to carry critical information for face perception (Costen et al., 1994, 1996; Gold et al., 1999; Hayes et al., 1986). It would be useful for a future study using methods like those in the current study to investigate the nature of developmental changes in the use of spatial frequency information in recognizing facial expressions before 10 years of age.

6. General discussion

In all three experiments, masking of mid spatial frequencies had the most impact on adults' contrast thresholds for recognition of both facial expressions and facial identity. This pattern is consistent with most previous studies (Costen et al., 1994, 1996; Fiorentini et al., 1983; Gold et al., 1999; Goren & Wilson, 2006; Näsänen, 1999; Ojanpää & Näsänen, 2003; Schwartz et al., 1998; Schyns & Oliva, 1999). The tuning was also largely scale invariant: as distance tripled in Experiment 2 from 60 to 180 cm, the critical spatial frequency band shifted only slightly lower for both tasks. These patterns are consistent with previous findings that adults process both identity and facial expression holistically (Calder, Young, Keane, & Dean, 2000; Maurer, Le Grand, & Mondloch, 2002) rather than merely as a collection of independently processed features that would be most easily discriminated with higher spatial frequencies. Such holistic processing would be expected to be tied to object size (e.g., face width) rather than to retinal size.

Despite the similarities in the results for facial identity and facial expression, there was a significant difference that replicated across the three groups of adults tested at 60 cm in Experiments 1, 2, and 3, and that persisted at 120 and 180 cm in Experiment 2: in every case, the critical spatial frequency band identified by the contrast thresholds peaked at a higher masked spatial frequency for facial expression than for facial identity. This pattern does not reflect the difference in the available information for distinguishing facial identity and facial expression, since an ideal observer behaved similarly for the two types of task. Instead, it is the result of (1) greater efficiency in processing the available information at low and mid spatial frequencies when performing either task, and (2) a drop of sensitivity at high spatial frequencies that begins at a lower value when judging identity than when judging expression. The non-overlapping peaks and elbows provide a new piece of evidence for the, at least partial, separation of the neural systems underlying the processing of facial identity and facial expression (Bruce & Young, 1986; Haxby et al., 2000). The difference indicates that more details are needed to recognize facial expression than to recognize facial identity. As a result, identity may be (slightly) easier to recognize than expression under conditions that degrade the transmission of higher spatial frequencies in a face image such as great distance and poor lighting.

As suggested by many researchers (de Gardelle & Kouider, 2010; Rotshtein, Vuilleumier, Winston, Driver, & Dolan, 2007), the recognition of facial expressions involves two pathways: a fast subcortical pathway relying on low spatial frequency information that does not require awareness and a slower cortical pathway relying on higher spatial frequency information, and requiring awareness. It is likely that the current results reflect the activity of the cortical pathway, since the participants were performing a recognition task with high visual awareness.

Experiment 3 indicates that children at age 10 and 14 also rely on low and mid spatial frequencies to recognize facial identity and facial expression, with the elbows of the functions in a similar location to those of adults. The similar patterns suggest that, like adults, children rely on finer details to judge facial expression than to judge identity. These findings are consistent with evidence that children begin to process faces holistically by 4–6 years of age (de Heering, Houthuys, & Rossion, 2007; Mondloch, Pathman, Maurer, Le Grand, & de Schonen, 2007; Pellicano & Rhodes, 2003; Tanaka, Kay, Grinnell, Stansfield, & Szechter, 1998). However, children's ability to recognize facial identity and facial expression continues to develop after age 10 (Gao & Maurer, 2009; Mondloch et al., 2003, 2004, 2002). Such a long developmental course is consistent with the results of the current study. Even at age 14, children needed more contrast than adults to recognize both facial identity and facial expression. The developmental difference may not be face-specific but arise from general differences in extracting signals from noise, selective attention, and/or high-level visual abilities (e.g., Kovács et al., 1999; Mondloch, Maurer, & Ahola, 2006). It may also be related to the fact that adolescents' brain areas involved in the processing of facial identity and facial expression are not yet as specialized as in adults (Golarai, Liberman, Yoon, & Grill-Spector, 2010; Lobaugh, Gibson, & Taylor, 2006).

The current results suggest that in everyday interactions, under poor lighting (low contrast), children may be especially poor at recognizing facial identity and facial expression compared to adults. However, we used static images of faces in the current study. It is possible that in real life, dynamic information in faces may help children to recognize facial identity and facial expressions and eliminate the deficit compared to adults. In the current study, we also used only adults' faces. Although children see adults' faces in their everyday lives, they may see more faces of age mates or find them more salient. Indeed, there is some evidence for an own age advantage in the recognition of facial identity (Anastasi & Rhodes, 2005) that may reflect differential spatial frequency tuning. Future studies could use methods similar to the current study to investigate the development of spatial frequency tuning with children's and dynamic faces, as well as test children at other ages not included in the current study.

Acknowledgment

The research was supported by a grant from the Canadian Natural Sciences and Engineering Research Council (9797) to DM. We thank Patrick J. Bennett for help creating the stimuli and Shirley Ma for assistance in collecting the data.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2011.01.011.

References

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin and Review, 12*, 1043–1047.
Bachmann, T. (1991). Identification of spatially quantised tachistoscopic images of faces: How many pixels does it take to carry identity? *European Journal of Cognitive Psychology, 3*, 87–103.
Benedek, G., Benedek, K., Kéri, S., & Janáky, M. (2003). The scotopic low-frequency spatial contrast sensitivity develops in children between 5 and 14 years. *Neuroscience Letter, 345*, 161–164.
Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–439.
Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305–327.
Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 527–551.
Conrey, B., & Gold, J. (2009). Pattern recognition in correlated and uncorrelated noise. *Journal of the Optical Society of America A, 26*, B94–B109.
Costen, N. P., Parker, D. M., & Craw, I. (1994). Spatial content and spatial quantisation effects in face recognition. *Perception, 23*, 129–146.
Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception and Psychophysics, 58*, 602–612.

de Gardelle, V., & Kouider, S. (2010). How spatial frequencies and visual awareness interact during face processing. *Psychological Science, 21*, 58–66.

de Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at 4 years of age: Evidence from the composite face effect. *Journal of Experimental Child Psychology, 96*, 57–70.

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.

Ellemberg, D., Lewis, T. L., Maurer, D., Liu, C., & Brent, H. P. (1999). Spatial and temporal vision in patients treated for bilateral congenital cataracts. *Vision Research, 39*, 3480–3489.

Etcoff, N. L. (1984). Selective attention to facial identity and facial emotion. *Neuropsychologia, 22*, 281–295.

Fiorentini, A., Maffei, L., & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception, 12*, 195–201.

Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology, 102*, 503–521.

Gao, X., Maurer, D., & Nishimura, M. (2010). Similarities and differences in the perceptual structure of facial expressions of children and adults. *Journal of Experimental Child Psychology, 105*, 98–115.

George, M. S., Ketter, T. A., Gill, D. S., Haxby, J. V., Ungerleider, L. G., Herscovitch, P., et al. (1993). Brain regions involved in recognizing facial emotion or identity: An oxygen-15 PET study. *Journal of Neuropsychiatry and Clinical Neurosciences, 5*, 384–394.

Golarai, G., Liberman, A., Yoon, J. M. D., & Grill-Spector, K. (2010). Differential development of the ventral visual cortex extends through adolescence. *Frontiers in Human Neuroscience, 3*, 80. doi:10.3389/nero.09.080.2009.

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

Goren, D., & Wilson, H. R. (2006). Quantifying facial expression recognition across viewing conditions. *Vision Research, 46*, 1253–1262.

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

Hadad, B., Maurer, D., & Lewis, T. L. (2010). The effects of spatial proximity and collinearity on contour integration in adults and children. *Vision Research, 50*, 772–778.

Harmon, L. D. (1973). The recognition of faces. *Scientific American, 229*, 71–82.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32*, 203–218.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223–232.

Hayes, T., Morrone, M. C., & Burr, D. C. (1986). Recognition of positive and negative bandpass-filtered images. *Perception, 15*, 595–602.

Hornak, J., Rolls, E. T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia, 34*, 247–261.

Kovács, I., Kozma, P., Fehér, Á., & Benedek, G. (1999). Late maturation of visual spatial integration in humans. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 12204–12209.

Leonard, H. C., Karmiloff-Smith, A., & Johnson, M. H. (2010). The development of spatial frequency biases in face recognition. *Journal of Experimental Child Psychology, 106*, 193–207.

Lobaugh, N. J., Gibson, E., & Taylor, M. J. (2006). Children recruit distinct neural systems for implicit emotional face processing. *Neuroreport, 17*, 215–219.

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

Mondloch, C. J., Dobson, K. S., Parsons, J., & Maurer, D. (2004). Why 8-year-olds cannot tell the difference between Steve Martin and Maul Newman: Factors contributing to the slow development of sensitivity to the spacing of facial features. *Journal of Experimental Child Psychology, 89*, 159–181.

Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology, 86*, 67–84.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*, 553–566.

Mondloch, C. J., Maurer, D., & Ahola, S. (2006). Becoming a face expert. *Psychological Science, 17*, 930–934.

Mondloch, C. J., Pathman, T., Maurer, D., Le Grand, R., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence of adult-like holistic face processing. *Visual Cognition, 15*, 564–577.

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

Ojanpää, H., & Näsänen, R. (2003). Utilisation of spatial frequency information in face search. *Vision Research, 43*, 2505–2515.

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods Instruments and Computers, 36*, 634–638.

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

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science, 14*, 618–622.

Rotshtein, P., Vuilleumier, P., Winston, J., Driver, J., & Dolan, R. (2007). Distinct and convergent visual processing of high and low spatial frequency information in faces. *Cerebral Cortex, 17*, 2713–2724.

Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research, 70*, 273–292.

Schwartz, O., Bayer, H., & Pelli, D. (1998). Features, frequencies, and facial expressions (Abstract). *Investigative Ophthalmology and Visual Science, 39*, 173.

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.

Sergent, J., Ohta, S., Macdonald, B., & Zuck, E. (1994). Segregated processing of facial identity and emotion in the human brain: A PET study. *Visual Cognition, 1*, 349–369.

Smith, F. W., & Schyns, P. G. (2009). Smile through your fear and sadness: Transmitting and identifying facial expression signals over a range of viewing distances. *Psychological Science, 20*, 1202–1208.

Tanaka, J. W., Kay, J. B., Grinnell, E., Stansfield, B., & Szechter, L. (1998). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition, 5*, 479–496.

Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research, 35*, 3053–3069.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*, 242–249.

Tranel, D., Damasio, A. R., & Damasio, H. (1988). Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology, 38*, 690–696.

Winston, J. S., Henson, R. N. A., Fine-Goulden, M. R., & Dolan, R. J. (2004). FMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology, 92*, 1830–1839.

Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological Research, 48*, 63–68.

Young, A. W., Newcombe, F., de Haan, E. H., Small, M., & Hay, D. C. (1993). Face perception after brain injury: Selective impairments affecting identity and expression. *Brain, 116*, 941–959.