Psicológica (2014), 35, 171-194.

A comparison of backward masking of faces in expression and gender identification

Luis Aguado*, Ignacio Serrano-Pedraza, and Ana García-Gutiérrez

Universidad Complutense, Madrid, Spain

The effects of different masking conditions on identification of face gender and expression at different target durations (17-119 ms time range) were studied in two experiments. In Experiments 1a and 1b, the effects of face masks were compared against those of noise masks (scrambled face stimuli) and a control, no-mask condition. Significant masking by noise masks was only found at the shortest target duration (17 ms). Effective masking by face masks was observed over a slightly longer time window in the gender than in the expression task (17-85 and 17-51 ms respectively). Moreover, clearly differentiated effects of face and noise masks were observed only in the expression task. In Experiments 2a and 2b, faces quantized with different sampling sizes were used as masks. A graded effect of sampling size was observed in the expression task, with those masks that preserved more facial information exerting stronger masking. However, all masks were equally effective in the gender task. These results demonstrate an interaction between masking and task demands, suggesting that different processing mechanisms may underlie identification of different properties of faces. An interpretation is offered in terms of the relative role of configural and feature processing in expression and gender identification.

The ability to process information from faces is extremely important for humans because faces provide crucial information about different person properties of high relevance for social interaction. It is known that face processing takes place very fast and that accurate discrimination of faces can be performed rapidly (e.g., Lehky, 2000; Pegna, Khateb, Michel, & Landis, 2004; see Palermo & Rhodes, 2007, for a review). A special case is that facial expressions of emotion can be accurately identified even when

* This work was supported by grant Psi2010-18682 from the Spanish Ministerio de Ciencia e Innovación. Correspondence concerning this article should be addressed to Luis Aguado, Facultad de Psicología, Universidad Complutense, Campus de Somosaguas, 28223, Madrid, Spain. E-mail: laguado@ucm.es

the face is presented under restricted viewing conditions due to short exposure time and/or masking. For example, emotional faces have a lower perceptual threshold than non-emotional faces (Calvo & Esteves, 2005), can be identified with high accuracy with short presentation durations (e.g., Calvo & Lundqvist, 2008) and are quickly detected amongst distracters (e.g., Öhman, Lundqvist, & Esteves, 2001). Results such as these have led to the proposal that emotional expressions are especially salient and effective in commanding attention (Calvo & Nummenmaa, 2008).

Visual identification of objects such as faces is a gradual process that develops over a short period of time. This has been related to the time taken to complete the feed-forward sweeps of neural activity initiated by the sensory input and to the number of iterations or feedback sweeps between the cortical regions involved in the analysis of the stimulus (Enns, 2004; Enns & DiLollo, 2000). Theories of face perception (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000) and empirical evidence indicate that different face processing tasks, such as identifying the expression or the gender of a face, may rely on different types or amounts of visual information (Atkinson, Tipples, Burt, & Young, 2005; Le Gal & Bruce, 2002; Schyns & Oliva, 1999; Schweinberger & Soukoup, 1998) and be based on different underlying neural mechanisms and brain systems (Dzhelyova, Ellison, & Atkinson, 2011; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000).

Backward masking provides a convenient method to study the dynamics of information processing by the visual system (see Breitmeyer & Ogmen, 2000, for a review). In this paradigm, perception of a briefly presented stimulus is impaired when it is immediately followed by another visual stimulus that acts as a mask. In research on facial expressions of emotion, backward masking has been used as a tool to study specific issues such as the duration thresholds for accurate identification of different expressions (Milders, Sahraie, & Logan, 2008; Roesch, Sander, Mumenthaler, Kerzel, & Scherer, 2010) or the possibility that emotional faces can control behavioral and physiological reactions even when they are processed outside of awareness (Bunce, Bernat, Wong, & Shevrin, 1999; Lidell, Williams, Rathjen, Shevrin, & Gordon, 2004; Murphy & Zajonc, 1993).

In the present series of experiments we used the backward masking paradigm with different target durations and mask types. These manipulations were aimed at exploring two related issues. First was the exposure time needed for expression and gender identification. Exposure time is a crucial variable that directly determines the amount of information that reaches the visual system and backward masking allows a precise control of the duration of effective exposure to the target stimulus. Given that backward masking is thought to act by interrupting processing of the target stimulus, studying the effects of different exposure times on gender and expression identification could help us to infer the time needed to extract from the stimulus the information required for accurate performance on these tasks.

The second issue to explore was the influence of the facial information content of the mask on expression and gender identification. Studies that have compared the effects of facial and non-facial masks (that is, masks that do not contain face stimuli) or of masks that contain different levels of facial information, have shown that the presence of facial information in the mask is crucial to produce significant masking in face identification (Costen, Shepherd, Ellis, & Craw, 1994; Loffler, Gordon, Wilkinson, Goren, & Wilson, 2005). This result indicates that masking acts by interfering specifically with face processing mechanisms. However, the extent to which this conclusion can be applied to different categorization tasks is not known because the effects of backward masking in different tasks have not been compared. Given that masking is highly dependent on the precise information content of the mask this comparison might reveal between task variations in the sensitivity to interference by masks with different visual content. This is theoretically relevant because these variations would suggest that each task is mediated by different perceptual mechanisms.

Comparing the effects of target duration and mask type on the identification of gender and expression can provide relevant information in relation to three aspects in which identification of facial expressions of emotion have been said to differ from other face recognition tasks. First, it has been frequently claimed that emotionally expressive faces and specially those conveying threat are quickly processed and recognized based on simple stimulus features and schematic processing (e.g., Öhman, 1993; Öhman & Mineka, 2001). No similar claim has been made in relation to gender identification and though accurate performance in this task has been reported with exposure times over 50 ms (Roesch et al., 2010), a direct comparison of the effects of different masks and exposure time conditions on expression and gender identification is yet to be made.

Second, from an adaptive point of view it seems reasonable to assume that evolutionary pressures to evolve mechanisms for fast detection based on simple or incomplete input should have been stronger for emotional expression than for gender. This is not to deny that gender also has high relevance in social encounters. But while gender is a stable property of faces and thus may be crucial when we encounter a new person, facial expression changes rapidly within the same individual and signals specific actions trends and dispositions on a moment to moment basis, demanding in return fast and flexible reactions from the observer (Blair, 2003; Ekman, 1997). The distinction between the processing of fixed and variable properties may be a crucial characteristic of the visual processing of faces and it has been incorporated into modern brain-systems theories of face perception (Calder & Young, 2005; Haxby et al., 2000). According to this distinction, optimal detection of expression changes should be especially dependent on fast perceptual processes. A finding consistent with these assumptions is that neural activity from different brain regions can be evoked by masked emotional faces at short latencies after stimulus onset (e.g., Whalen, Rauch, Etcoff, McInerney, Lee, & Jenike, 1998; for reviews see Adolphs, 2006; Vuilleumier & Pourtois, 2007). More specifically, event-related potential (ERP) studies have revealed that the N170 component, detected over occipito-temporal sites with peak amplitudes around 170 ms after stimulus onset, is modulated by facial expression but not by face gender (e.g., Ashley,Vuilleumier, & Swick, 2004; Mouchetant-Rostaing et al., 2000). Given that the N170 is the first visual ERP component differentiating between faces and other types of objects (e.g., Bentin, Allison, Puce, Perez, & McCarthy, 1996; Bötzel, Ecker, Mayer, Schulze, & Straube, 1995), its differential sensitivity to expression and gender suggests faster identification of the emotional expression of faces.

Third, studying the sensitivity of expression and gender identification to mask type and exposure time is also relevant for theories of face perception that propose independent processing of different facial properties (Bruce & Young, 1986) or that assume that stable and changing characteristics of faces (e.g., gender vs. expression) are processed by different neural systems (Calder & Young, 2005; Haxby et al., 2000). If tasks differ in the relative diagnostic value of different facial features or in the specific processing demands required in each case, this might be reflected in their sensitivity to stimulus duration and masking, due to differences in the type and amount of information needed for accurate identification or in the time needed to perform the computations required by each task.

In the experiments reported here, participants performed expression or gender identification tasks under different conditions of target duration and mask type. These manipulations influenced both the visibility of the target and the time available to extract and process the information required to perform the task. In Experiments 1a and 1b the masking power of neutral faces and noise masks (scrambled face stimuli) was compared. In Experiments 2a and 2b, the information content of masks was manipulated by using quantized neutral faces with different sampling sizes that preserved different levels of facial information. Although variation of target duration in backward masking designs has been used frequently in studies on perception of facial expression, the effects of mask type have rarely been explored and this only in tasks that required familiarity judgments or individual face recognition (Costen et al., 1994; Loffler et al., 2005). Differences in the influence of target duration in the gender and expression tasks, for example accurate identification of expression with shorter stimulus durations, would suggest that these two tasks differ in terms of the time needed to extract relevant information from the face. Second, differences in the effects of mask content would be suggestive of differences in the perceptual processes underlying the identification of the gender and expression of faces, for example the relative role of configural or feature-based processing in each task.

EXPERIMENTS 1A AND 1B

The following two experiments were aimed at evaluating the effects of backward masking on the discrimination of the expression and gender of faces. In these experiments, the effect of a face mask was compared against that of a visual noise mask and a non-mask control condition. Face targets were presented at five different exposure times, while mask duration was kept constant. This design can provide crucial information to adequately compare and interpret the effects of backward masking on the identification of the expression and the gender of faces. Comparing face masks with visual noise masks that preserve similar low-level physical properties but do not contain facial information will provide relevant evidence on the extent to which masking is dependent on the presence of this information in the mask and is thus due to specific interference of face processing mechanisms.

Male and female faces showing happy, angry or neutral expressions were used as targets (see examples in Figure 1a). Neutral faces were included in order to avoid the possibility that in the expression task participants might adopt decision rules that would preclude explicit identification of all expressions (for example, categorizing as “angry” all faces that do not show a smile). Sampled neutral faces of different identities to those of the targets were used as face masks (each sample was a square of 8 x 8 pixels and it was quantized using the average density of the original image; see Figure 1a). Noise masks were those neutral faces but with the samples randomly reordered so that no facial information was preserved. Target faces were presented for varying durations (17, 34, 51, 85 and 119 ms) that were chosen based on previous research on emotion discrimination with backward masked faces (e.g., Esteves & Öhman, 1993; Maxwell & Davidson, 2004; Pessoa, Jappee, & Ungerleider, 2005). We expected shorter durations to provide the more sensitive condition to reveal differences between expression identification and gender identification. Different participants performed the expression and gender tasks. In the expression task the participants had to identify the expression of the face (happy, angry or neutral), while in the gender task they had to identify the face as masculine or feminine.

METHOD

Participants. Participants were twenty four psychology students (M_{age} = 21.59; range 18-27) who participated in experiments for course credit. Twelve participants were assigned to each task (11 women and 1 men to the expression task and 12 women to the gender task).

Apparatus and Stimuli. Presentation of stimuli and registration of responses was controlled through the E-Prime 2.0 software. Stimuli were presented on a VGA 17-in. monitor (refresh rate 59.82 Hz). Participants were seated at a distance of 50 cm from the screen with head position controlled with a head and chin rest. Responses were registered through a five-key response box (PST Serial Response Box, 200A). Sessions were carried out individually in a soundproof, dimly lit room. Stimuli were 48 pictures of human male and female showing a happy, a neutral or an angry expression, taken from the Karolinska Directed Emotional Faces (KDEF) collection (Lundqvist, Flykt, & Öhman, 1998). Each facial expression was presented the same number of times. In order to avoid possible influences of hairstyling and promote attention to internal facial features images were cut to conceal most of the hair. Moreover, with the aim of reducing the saliency of the teeth white due to the opening of the mouth in happy and angry faces and preventing the participants from basing their responses exclusively on this feature, the mouth area was blurred using a Gaussian filter of 10 pixel (see Figure 1a). Images were equated in contrast energy (root mean square contrast or cRMS = 0.2) following the same method used before in our laboratory (Aguado, García-Gutiérrez, & Serrano-Pedraza, 2009). Eight additional neutral faces from the KDEF collection (four male, four female) that were processed so that pixels were visible were used as masks.

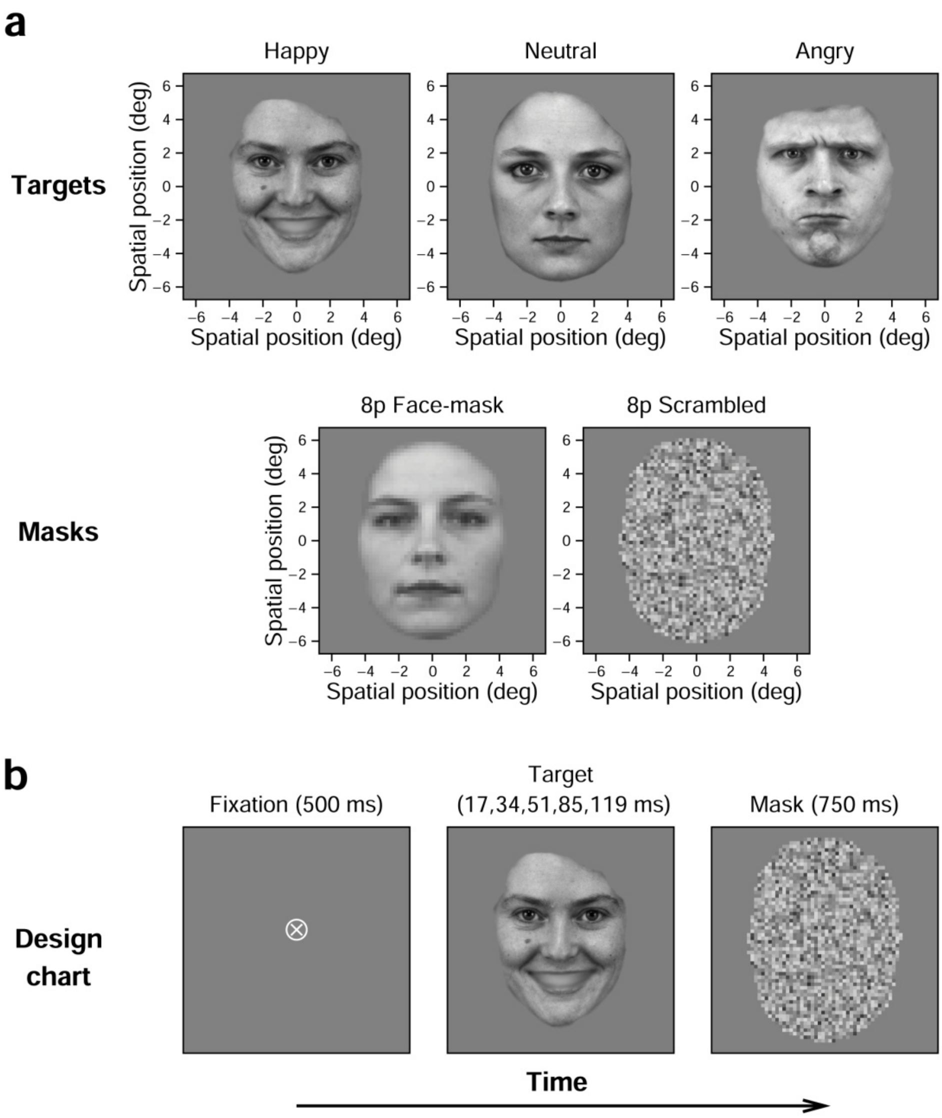

Figure 1. Example of stimuli used in the experiments. (a) Examples of targets with different emotional expressions and examples of masks. The mask on the left is a sampled and quantized neutral face and the size of the sampling is 8 x 8 pixels each sample was quantized using the average density of the original image. The mask on the right is the same image presented on the left but with samples scrambled randomly. (b) Design chart showing the temporal course of one trial of backward masking in the noise-mask condition.

Procedure. On experimental trials (see Figure 1b), target faces were presented on the center of the screen. The faces subtended an angle of 6 x 10 degrees (horizontal x vertical) of visual angle (deg). They were presented inside of a grey square (45.2 cd/m^2) with a size of 512 x 512 pixels, subtending an area of 13.5 x 13.5 degrees. Five exposure times were used (17, 34, 51, 85 and 119 ms). Longer durations were not used because we observed in a pilot experiment that performance reached asymptotic level at durations over 120 ms. Each trial was preceded by a 500 ms fixation mark centered on the screen. Three different conditions were included: face-mask, noise-mask and no mask (see Figure 1b). In the face-mask and noise-mask conditions the target, presented at each of the five exposure times, was immediately followed by the corresponding mask, that was terminated by the participant´s response or stayed on for a maximum of 750 ms. In the no mask condition, the target was followed by a screen containing an empty grey square with similar dimensions to the background on which the faces were presented. This blank period was terminated by the participant´s response or lasted for a maximum of 750 ms. In masking trials, masks were programmed so that each target face was followed an equal number of times by a mask of the same or the other gender. In order to prevent anticipation of the next trial, a varying inter-trial interval was used (M = 2212 ms; range 1500-3000 ms).

The experimental task consisted of identifying the expression (Experiment 1a) or the gender (Experiment 1b) of the target face. In the expression condition, participants had to identify the target faces as showing a happy, angry, or neutral expression by pressing keys 1, 3, and 5 of the response box. The central key (key 3) was always assigned to the "neutral" response, while assignment of the 1 and 5 keys to the "happy" or "angry" response was counterbalanced. For the Experiment 1b (gender), participants had to press keys 1 or 5, for female or male, counterbalanced. In both experiments, response's options appeared written in a line of the screen, and then the participant was forced to choose one of the options. In order to avoid interference of responses effects, options on the screen and the corresponding key were localized in the same place. There were a total of 720 trials (144 trials x 5 durations). Before starting the experimental phase, twenty practice trials were run. To this end were used different target faces that had not been shown during the rest of the experiment.

RESULTS

Expression task

Figure 2 (left panel) shows the proportion of correct identifications of expressions as a function of the target exposure times and for the three masking conditions: face-mask (samples of 8 x 8 pixels), no-mask, and noise-mask (scrambled samples of the face masks). As can be seen in the figure, performance in the expression task reached a high level of accuracy at all target durations in the no-mask control condition. The poorest performance was obtained in the face-mask condition, with strong masking effects appearing over the 17-51 ms time window and accuracy increasing as a positive function of exposure time. On the other hand, the noise-mask condition only had an effect at the shortest 17 ms exposure time. Performance under this condition was comparable to the no-mask control at all other target durations.

For all repeated measures ANOVA analyses reported in the present paper, the Greenhouse-Geisser correction was applied when the sphericity assumption was violated. Post-hoc analyses were performed using the Bonferroni correction (significant when $p \leq .05$). A 5 (Duration) x 3 (Mask) repeated measures ANOVA with Target Duration and Masking condition as factors gave significant main effects of Duration, F (4, 44) = 36.02, (MSe = .004), $\eta_p^2 = .76$, Mask, F (2, 22) = 61.21, (MSe = .005), $\eta_p^2 = .85$ and the interaction, F(8, 88) = 26.73, (MSe = .002), $\eta_p^2 = .71$. Post-hoc analyses performed to explore this interaction revealed significant masking effects in the noise-mask condition only at the shortest target duration of 17 ms ($p_s \leq .001$). A stronger effect of the face-mask than of the noise-mask was also observed at this short exposure time. Significant masking effects of the face mask were also observed at the 34 and 51 ms durations ($p_s \leq .01$). A graded effect of exposure time, with significant increases in accuracy as target duration increased, was obtained only in the face-mask condition. In the noise-mask condition the only significant differences were between 17 ms and the other durations. Finally, in the no mask condition the duration of the target had no effect whatsoever.

Gender task

Figure 2 (right panel) shows the results of the gender task under all three masking conditions. It can be seen that the highest and lowest accuracy levels corresponded to the no mask and the face-mask condition respectively, with intermediate accuracy in the noise-mask condition. Moreover, although the effect appears smaller in the no-mask condition, an overall influence of exposure time can be seen, with increasing durations leading to increased accuracy. These impressions were confirmed by statistical analysis. Significant main effects were obtained of Target Duration, F (4, 44) = 36.9, (MSe = .004), η_p^2= 0.77; Mask, F(2, 22) = 39.2, (MSe = .004), η_p^2= .78 and the interaction, F(8, 88) = 6.7, (MSe = .003), η_p^2= .38. Post-hoc comparisons yielded significant masking effects in the face-mask condition at 17, 34, 51and 85ms. At the longer 119 ms target duration, a marginally significant masking effect (p =.09) was still obtained in this condition. However, in the noise-mask condition significant masking was only observed at the shortest 17 ms exposure time (p_s < .003). As to the comparison between the face and the noise masks, significant differences were not found at any target duration, although marginally significant effects appeared at the 17, 34, and 51ms durations (significance levels between .06 and .08). Finally, significant effects of exposure time were found at all three masking conditions.

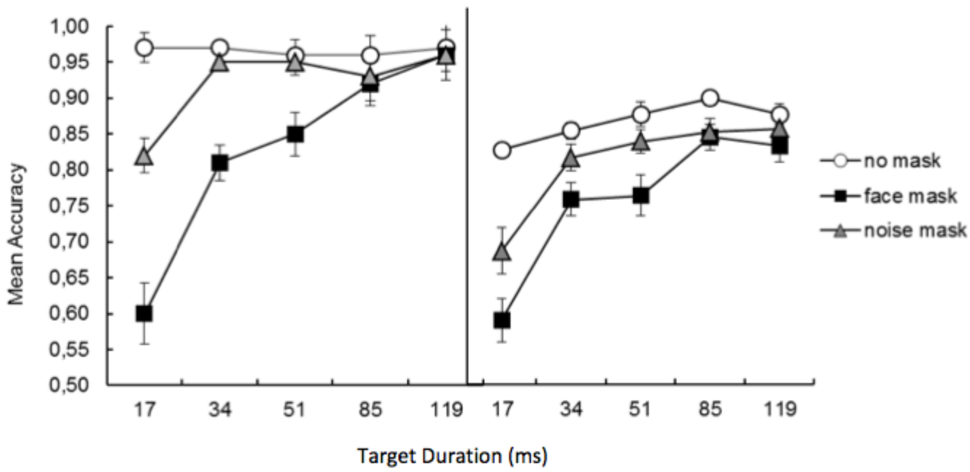

Figure 2. Mean accuracy in Expression task (Experiment 1a-left panel) and Gender task (Experiment 1b-right panel). Error bars represent SEM.

DISCUSSION

Effects of backward masking on expression and gender identification were observed in Experiments 1a and 1b at different target durations. Maximal accuracy in the expression task was found at all exposure times in the no mask condition, showing that in the absence of masking the emotional expression of very briefly presented faces can be accurately identified (see also Calvo & Lundqvist, 2008). Overall, less accurate performance was observed in the gender task. Although performance in these two tasks cannot be directly compared due to the different number of response alternatives (three in the expression task but only two in the gender task), the fact that even under these conditions superior performance was observed in the expression task indicates that identifying the gender of a face was a relatively more difficult task.

Backward masking, defined as an accuracy level significantly lower than in the no mask condition, was observed over the 17-51 ms time window in the expression task, with stronger effects at the shortest target durations. However, identification of facial expression reached above chance levels even at these short durations. No evidence of masking was obtained in this task at the two longest durations of 85 and 119 ms. This indicates that information that allows identification of emotional expression with maximal accuracy can be extracted from a face with 85 ms exposure, at least in the case of happy and angry expressions. The time window for effective backward masking was slightly longer in the gender task, with significant masking at 85 ms and a non significant trend still apparent at the longest duration of 119 ms. These results suggest that longer exposure to the stimulus is needed to identify the gender of faces, probably because this task requires more complete or detailed information than what is needed for expression identification. As to the effect of mask type, significant masking effects beyond the shortest target duration of 17 ms were found only when the targets were followed by face masks. Thus, backward masking in both tasks depended on the presence of facial information in the mask.

The effective masking produced by the noise mask with the shortest target duration suggests that interference of visual processing can be produced at very early stages, when the low-level properties of the face image are being processed. However, it has to be pointed out that a differential effect of mask type was still observed at this target duration. This is interesting because it suggests that some interference with mechanisms specific to the visual coding of faces might take place even at these early stages of perceptual processing. On the other hand, the fact that beyond 17 ms significant masking was only produced by the face masks suggests that masking at these longer target durations only occurs if processing of the mask interferes with mechanisms specific to face processing. However, the results from the gender task are somewhat ambiguous. Although it is true that in this case significant masking beyond the 17 ms duration was only produced by the face masks, the noise mask had an effect that was not significantly different from that of either the face mask or the no-mask condition. Consequently, the results of Experiment 1b do not allow us to draw firm conclusions regarding the role of facial information from the mask on masking in the gender task. With the aim of evaluating more precisely the extent to which the presence of facial information in the mask is critical to produce backward masking in the gender and expression tasks, the effect of face masks that preserved different levels of facial information was compared in Experiments 2a and 2b.

EXPERIMENTS 2A AND 2B

A more sensitive method to solve the question of the relative influence of facial information in the mask in the gender and expression tasks might be to compare the effect of masks that are not equally recognizable as faces. It is well known that face recognition is impaired by sampling and quantizing faces (Harmon & Julesz, 1973). Quantization introduces noise that interferes with information relevant for facial recognition. Because of this, identification accuracy of pixelated faces varies as a direct function of quantization levels (e.g., Bachman, 1991). In the following experiments we used face masks with different sampling sizes that preserved different degrees of facial information (see examples on the first column of Figure 3). If masking interferes specifically with face processing mechanisms, masks that preserve more facial information should be more interfering and thus should exert a stronger masking effect. A differential effect of mask type on expression and gender identification would suggest differences in the perceptual mechanisms underlying each of these tasks.

METHOD

Participants. Participants were twenty four psychology students (M_{age} = 20.08, range 18-24) who participated in the experiments for course credit. Twelve participants were assigned to each task (7 women and 5 men to the expression task and 11 women and 1 man to the gender task).

Procedure. The apparatus, stimuli and experimental procedure were similar to those of Experiments 1a and 1b. The exception was that the noise mask was not used and that each face mask was presented with three different sampling sizes. Each sample was quantized using the average density of the original images and three different versions of each mask were prepared. Distortion of the image and consequently the difficulty to recognize it as a face is higher the bigger the size of the squares. In the present experiments we used masks with three different sampling sizes (squares of 8×8, 16×16, and 32×32 pixels).

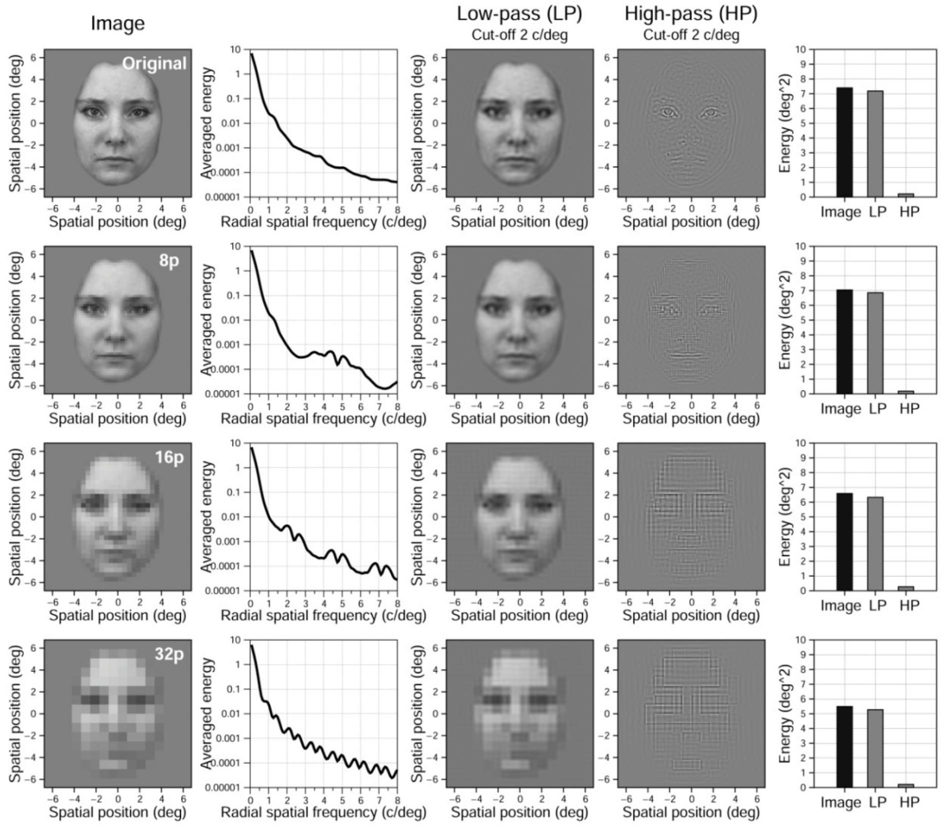

Figure 3. Spectral analysis of face-mask stimuli used in the experiments. First column, a-d: examples of one original image and the corresponding quantized versions, with sampling sizes of 8 x 8 pixels (8p), 16 x 16 pixels (16p) and 32 x 32 pixels (32p). The second column shows the averaged energy of the images as a function of the radial spatial frequency (see Sierra-Vazquez & Serrano-Pedraza (2010, pp.788). The third and fourth columns show the images filtered low-pass (LP) and high-pass (HP) respectively. The filters used were isotropic ideal filters with cut-off frequency of 2 c/deg for both LP and HP. The fifth column shows the energy in deg^2 of the unfiltered image and the energy of both LP and HP images. Note that the contrast of the HP images has been increased artificially in order to make them easily visible.

RESULTS

Expression task

Figure 4 (left panel) shows the proportion of correct responses for expression identification as a function of Target Duration and Mask Type. As can be observed in the figure, identification accuracy increased with increasing exposure times under all three mask conditions. A clear effect of mask type is also observed, with more effective masking being produced by masks with higher sampling sizes. Stronger masking and thus less accurate identification of emotional expression was produced by the 8p masks, followed by the 16p and 32p masks. However, it can be seen that accuracy was well above chance level (33%) even at the shortest exposure time and with the more effective mask (8p), showing that discrimination of facial expression can be performed with relatively good accuracy even under very restricted viewing conditions. These impressions were confirmed by statistical analysis. A repeated measures ANOVA with a 3 (Mask Type) x 5 (Target Duration) design yielded significant main effects of Mask, $F(2, 22) = 29.22$ ($MSe = .003$), $\eta_p^2 = .72$ and Duration, $F(4, 44) = 17.16$ ($MSe = .010$), $\eta_p^2 = .60$. The interaction was also significant, $F(8, 88) = 8.23$, ($MSe = .003$), $\eta_p^2 = .42$. Post-hoc comparisons showed significant differential mask effects only at the two shortest target durations (17 and 34ms), with the poorest performance corresponding to the 8p mask condition. As to the effects of target duration, they were stronger under the 8 mask condition, with increasing target durations producing a steady and significant increase in accuracy. This effect was smaller with the 16p mask and was considerably reduced under the 32p mask condition, where the only significant difference was between the shortest and longest durations.

Gender task

As can be seen in Figure 4 (right panel), accuracy of gender discrimination improved with increasing exposure times under all three mask conditions. Statistical analysis confirmed these impressions, yielding only a significant main effect of Target Duration, $F(4, 44) = 49.2$, ($MSe = .005$), $\eta_p^2 = .81$. Paired comparisons between different stimulus durations were significant except for the 34 vs 51 and 85 vs 119 comparisons.

As stated in the procedure section, targets and masks included faces of both genders and care was taken to ensure that there was an equal number of trials where the target and mask faces were of the same or of the different gender. Thus, in the gender task half of the trials were congruent and the other half incongruent in terms of target and mask gender. It might be that identifying the gender of the target would have been more difficult on incongruent trials or that congruency modulated the effects of duration and masking. If this were so, it would compromise the generality of the analysis just reported. With the aim of evaluating this possibility, an additional analysis was performed including as a new factor the congruency between the gender of the target and mask faces. This analysis did not show a significant effect of Congruency, $F(1,11) = 3.7$, $p < .05$, or of any of the interactions involving this factor.

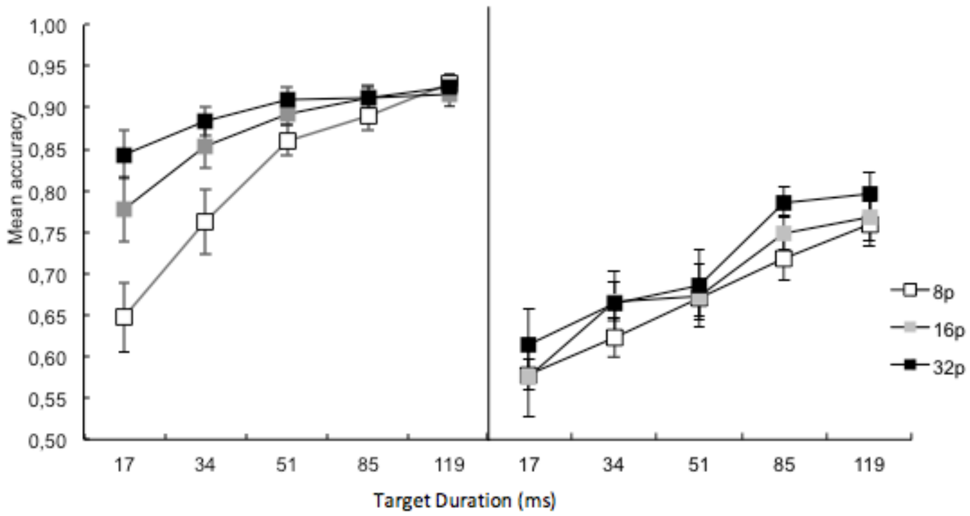

Figure 4. Mean accuracy in Expression task (Experiment 2a-left panel) and Gender task (Experiment 2b-right panel). Error bars represent SEM.

DISCUSSION

Similarly to Experiments 1a and 1b, a positive relationship was observed between target duration and accuracy, with improved performance in both tasks as duration increased. Overall, accuracy was also higher in the expression task, reaching above chance levels even under the more restricted viewing conditions. The most important result obtained in Experiments 2a and 2b was the differential influence of mask type on expression and gender discrimination. In the expression task, a graded effect of sampling size was observed and stronger masking was produced by the masks that preserved most facial information. However, in the gender task all three masks had similar effects.

The graded effect of mask type observed in the expression task indicates that the masking power of masks depended on the extent to which they preserved facial information. The strength of masking was critically dependent on the degree of "face-ness" of the masks, with the stronger masking effect corresponding to the 8p mask that preserved most facial information and could be clearly recognized as a face. This result suggests that masks act by interfering the ongoing process of building the perceptual representation of the target face. This graded effect was completely absent in the gender task, where all three masks had a comparable effect. It seems then that effective masking in this task can be explained as mainly due to some common property of the different mask types. One possibility is the first-order configural face information present in the three types of masks, with the three blobs corresponding to the eyes and mouth properly arranged. However, in view of the absence of significant differences between the effects of the face and noise mask observed in Experiment 1b, a role for low level physical properties of the masks cannot be excluded. In any case, the present experiments provide clear and new evidence of differential effects of backward masking in different face identification tasks. Possible interpretations of this result and their implications for theories of face processing are discussed below.

GENERAL DISCUSSION

Using a backward masking procedure, the present series of experiments compared the effects of target duration and mask content on gender and expression identification. The experiments had two main objectives. One was defining the time window over which backward masking occurs in these facial tasks. The second objective was to study the extent to which masking in gender and expression identification depends on the presence of facial information in the mask and can thus be attributed to specific interference with face processing mechanisms.

Backward masking was observed in the expression task over a 17-51 ms time window (Experiment 1a). In the gender task (Experiment 1b), the time window for effective masking was slightly longer, with significant effects at 85 ms and a trend still apparent at the longest, 119 ms target duration. Significant masking by noise masks was produced in both tasks only at the shortest 17 ms duration. Based on the assumption that processing of the target face is interrupted by the immediately following mask (Enns &

DiLollo, 2000; Kolers, 1968), our results can be used to infer the time needed to extract information from a face in order to identify its gender or its expression. The results obtained in the expression task suggest that exposure times around 80 ms are sufficient to extract information from a face that allow identification of its expression with maximal accuracy. This is considerably shorter than the durations reported in previous studies on individual face recognition (Costen et al., 1994; Loffler et al., 2005), although it is consistent with reports of relative resistance of facial expressions to backward masking, with above chance recognition of masked emotional faces with exposure times as short as 17-20 ms (Maxwell & Davidson, 2004; Milders et al., 2008; Pessoa et al., 2005; Szczepanowski & Pessoa, 2007). In fact, we also obtained above chance performance in the expression task (Experiments 1a and 2a) even under the more strict viewing conditions (17 ms exposure time and 8p face mask). Our estimation of exposure times for accurate identification of facial expression is also consistent with studies that have reported modulation of cortical responses at 80 ms post-stimulus onset by different emotional expressions (Eger, Jednyak, Iwaki, & Skrandies, 2003) and by liked vs. disliked faces (Pizzagalli, Regard, & Lehman, 1999).

The results obtained in the gender task suggest that the exposure time required for accurate identification of gender is in the 80-120 ms time range and so that it is longer than the time required for identification of expression. We have to recognize, though, that more firm conclusions about the differences between the gender and expression tasks would require a direct between-task comparison. However, this comparison was not possible due to the different number of response choices (three in the expression task and two in the gender task) and this no doubt involves a limitation of our study. In any case, the time required for accurate gender identification in the present study was again shorter than the duration threshold reported in the only published study where the effects of backward masking on expression and gender identification have been explicitly compared (Esteves & Öhman, 1993, their Experiment 1). In that study, similar duration thresholds on the 100-150 ms range were obtained for identification of gender and expression when the target was masked by neutral faces. A shorter threshold of 50 ms for gender identification was obtained in a more recent study by Roesch et al. (2010), but in this case no comparison was made with expression identification. Moreover, the masks used in that study were scrambled face stimuli similar to the noise masks used in our Experiment 1 that are not maximally effective to produce masking of target faces. However, care must be taken when comparing our results with those of previous experiments differing in procedural details and stimulus presentation method.

The second important result of our study was the different sensitivity of the expression and gender tasks to interference produced by masks with different levels of facial information. In Experiments 1a and 1b, where face and noise masks were compared, clear-cut effects of mask type were observed in the expression task. In this case, effective masking beyond the shortest, 17 ms target duration was only produced by the face masks. In contrast, a more ambiguous result was obtained in the gender task. Although in this case the face mask also produced effective masking compared with the control condition, the effect of the noise masks did not differ significantly either from that of the no-mask or the face-mask condition. Even though no clear conclusions can be drawn from this result, it suggests at least that the masking power of the face and noise masks is not as clearly differentiated as in the expression task. Additionally, it can be pointed out that the similar effect of both types of masks in the gender condition speaks against the possibility that incongruency between the gender of the target and mask faces influenced performance in the gender task. Should this be the case, superior performance would have been observed with the noise masks, when there was no such incongruency.

The two masks used in Experiments 1a and 1b were fairly dissimilar, as only the face mask contained facial information. Masks that preserved different levels of facial information were used in Experiments 2a and 2b to study its effects on the identification of the expression and gender of the target faces. This was done using as masks pixelated faces with different sampling sizes, thus making them more or less easily recognizable as faces. In this case, clear differences were observed between the expression and the gender tasks. Significant effects of sampling size were only found in the expression task, with the masks that preserved more facial information (that is, those with smaller sampling size) producing stronger masking. In contrast with this, accuracy of gender identification was comparable with all mask types. Masks that preserved different levels of facial information were equally effective to produce significant masking in this task.

The results obtained in the expression task are consistent with those reported by Loffler et al. (2005) in a matching task with expressively neutral faces. These authors also showed graded effects of target-mask similarity, with the more similar masks exerting the stronger masking effect. Although not completely equivalent, our image processing procedures also produced masks that differed in the extent to which they could be recognized as faces and so it is conceptually similar to that used in

Loffler et al's study. The new finding from our Experiments 2a and 2b is that performance was sensitive to this manipulation when the participant had to identify the expression of face but not when she had to identify its gender. This finding provides new evidence that the precise effects of backward masking of faces depend on the interaction between the nature of the mask and the specific target property that has to be identified on each occasion. This is to our knowledge the first demonstration that backward masking of faces has differential effects in different face identification tasks. It seems that the accrual of relevant information and the early processing operations needed to identify the gender of a face are highly sensitive to masking by unspecific and poorly structured visual information, so that significant masking effects are produced by a wide range of masks with relative independence of their informational content. On the other hand, the mechanisms involved in processing faces for expression identification seems to be sensitive to the degree of structural facial information present in the mask in a way that leads to graded effects depending on the extent to which the mask preserves facial information.

The results of the present series of experiments indicated that backward masking in expression and gender identification was effectively produced by face masks. However, our results also suggest that processing of a target face in the service of expression identification is relatively more resistant to interference by face masks. Potent masking effects were obtained with masks that preserved most of facial information, but when this information was degraded by quantization the impact of the masks was significantly reduced. In contrast, the results obtained in the gender task suggest that processing of faces for gender identification is easily impaired, being sensitive to interference by masks that only preserve coarse facial information. This pattern of results points to differences in the perceptual mechanisms involved in the identification of gender and expression. Although the present results do not allow any conclusion as to what those differences might be, a tentative interpretation can be offered in terms of the relative role of configural and feature processing in these two tasks. The facial expressions associated with basic emotions such as joy or anger consist on different configurations of feature displacements and shape changes produced by the joint movement of the eyes and mouth (Ekman & Friesen, 1978) and there is evidence that holistic and configural processing have a strong and specific influence on identification of facial expression (Calder, Young, Keane, & Dean, 2000; Calder & Jansen, 2005; Duranda, Gallayb, Seigneurica, Robichonc, & Badouin, 2007). This is not to deny that configural processing also contributes to gender identification or that the identification of the expression and gender of faces are completely dissociable tasks. Experimental effects suggestive of configural processing have indeed been reported in gender discrimination (e.g., Badouin & Humphreys, 2006) and demonstrations of symmetrical interference between expression and gender reveal that some facial features and second-order relations between features have common diagnostic value for both tasks (Aguado et al., 2009). However, local features and other types of cues that are irrelevant for discriminating facial expression have also diagnostic value for gender identification (Bruce et al., 1993; Bruce & Langton, 1994). Moreover, the characteristic biases for different spatial frequency bands that have been observed in expression and gender identification (Schyns & Oliva, 1999) are consistent with a different influence of configural and feature processing in these tasks (for examples of the spatial-frequency content of the different masks used in Experiment 2 see Figure 3). If configural processing plays a more important role in expression identification and feature processing in gender identification, our results might be interpreted as showing that these two types of perceptual processing are differentially sensitive to backward masking. If it is assumed that configural processing is especially fast or efficient, possibly because it does not require detailed processing of the face, strong interference might only occur when the visual representation of the target is displaced by a representation of the mask that is rich in facial information. More detailed or slower feature-based processing of the target such as that needed for gender identification would instead be more effectively interrupted by different types of face masks as long as they contain the minimal information needed to recognize a stimulus as a face, that is, the presence of blobs corresponding to the eyes and mouth in the proper configuration (e.g., Johnson & Morton, 1991). We must again recognize that this is only a speculative interpretation and that with independence of what is their most appropriate explanation our results are relevant in that they clearly show an interaction of backward masking with the nature of the task assigned to the participant.

One last issue worth comment refers to the locus where the target and the mask interact to produce backward masking. The stronger masking produced by the face masks suggest an interaction at the level of the systems where faces are encoded, possibly at the fusiform face area, that has been identified in numerous studies as a critical component of the human neural system for face processing (Haxby et al., 2000). On the other hand, the fact that significant effects of masking by the noise mask in Experiments 1a and 1b were only found at the shortest 17ms target duration suggest that noise is only effective at early stages of processing previous to coding of faces. This result is similar to that reported by Loffler et al.

(2005) who also found significant masking by noise masks at a short (27 ms) exposure time. These authors also proposed a locus of interaction at the level of the face coding system to explain the differential masking power of upright faces and inverted or scrambled faces. Finally, the differential sensitivity of expression and gender discrimination to the presence of facial information in the mask leaves open the intriguing possibility that backward masking could influence face perception at different processing stages depending on the task at hand.

RESUMEN

Comparación del enmascaramiento hacia atrás de caras en tareas de identificación de expresión facial y género. Dos experimentos evaluaron los efectos que distintas condiciones de enmascaramiento ejercían sobre la identificación de la expresión facial y el género de caras presentadas en distintas duraciones (17-119 ms). En los Experimentos 1a y 1b, los efectos de las caras-máscara se compararon con otras máscaras-ruido (caras aleatorizadas) y una condición de control, no enmascaramiento. Se observó un enmascaramiento significativo en las máscaras-ruido sólo en la duración más corta (17ms). Y un enmascaramiento más efectivo con caras-máscara en una ventana temporal ligeramente superior en la tarea de género que en la de expresión (17-85 y 17-51 ms, respectivamente). Además, sólo en la tarea de expresión, se observaron efectos claramente diferenciados entre las máscaras de cara y de ruido. En los experimentos 2a y 2b, se usaron como máscaras caras pixeladas a distintos niveles (8, 16 y 32 pixels). De nuevo, sólo en la tarea de expresión, observamos un efecto gradual relacionado con el tipo de máscara. Las máscaras que contenían más información facial ejercían un efecto de enmascaramiento más potente. Por el contrario, en la tarea de género, las máscaras fueron igualmente efectivas. Estos resultados demuestran una interacción entre las máscaras y las demandas de la tarea, sugiriendo que diferentes mecanismos de procesamiento puede subyacer a la identificación de distintas propiedades faciales. Los resultados se interpretan de acuerdo al papel relativo del procesamiento configuracional y de rasgo en la identificación de la expresión y del género.

REFERENCES

Adolphs, R. (2006). Perception and Emotion: How we recognize facial expressions. *Current directions in psychological science, 15*, 222-226

Aguado, L., García-Gutiérrez, A., & Serrano-Pedraza, I. (2009). Symmetrical interaction of sex and expression in face classification tasks. *Attention, Perception & Psychophysics,71,* 9-25. doi: 10.3758/APP.71.1.9

Ashley, V., Vuilleumier, P., & Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *Neuroreport*, *15*, 211-216.

Atkinson, A. P., Tipples, J., Burt, D. M., & Young, A. W. (2005). Asymmetric interference between sex and emotion in face perception. *Perception & Psychophysics*, *67*, 1199-1213.

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*, *1*, 87-103.

Badouin, J. Y., & Humphreys, G. V. (2006). Configural information in gender categorisation. *Perception*, *35*, 531-540.

Blair, R. J. (2003). Facial expressions, their communicatory functions neuro-cognitive substrates. *Philosophical Transactions of the Royal Society London Biological Sciences, 358*, 561-572.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: a comparison, review, and update. *Perception & Psychophysics, 62*, 1572-1595.

Bötzel, K., Ecker, C., Mayer, M., Schulze, S., & Straube, A. (1995). Frontal component of the somatosensory evoked potential. *Human Brain Maping, 3*, 245-253.

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception, 23,* 803-822.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305-327.

Bruce, V., Burton, M. A., Hanna, E., Healey, P., Mason, O., Coombes, A., ... Linney, A. (1993). Sex discrimination: How do we tell the difference between male and female faces? *Perception 22*, 131-152.

Bunce, S., Bernat, E., Wong, P., & Shevrin, H. (1999). Further evidence for unconscious learning: preliminary support for the conditioning of facial EMG to subliminal stimuli. *Journal of Psychiatric Research*, *33*, 341-347.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*, 641-651.

Calder, A. J., & Jansen, J. (2005). Configural coding of facial expressions: The impact of inversion and photographic negative. *Visual Cognition, 12*, 495-518. doi: 10.1080/13506280444000418

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. doi: 10.1037/0096-1523.26.2.527

Calvo, M. G., & Esteves, F. (2005). Detection of emotional faces: low perceptual threshold and wide attentional span. *Visual Cognition, 12* (1), 13-27.

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: salient physical features guide effective visual search. *Journal of Experimental Psychology: General, 137*, 471-94.

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behaviour Research Methods,40,* 109-115. doi: 10.3758/BRM.40.1.109

Costen, N. P., Shepherd, J. W., Ellis, H. D., & Craw, I. (1994). Masking of faces by facial and non-facial stimuli. *Visual Cognition, 1*, 227-251.

Dzhelyova, M. P., Ellison, A., & Atkinson, A. P. (2011). Event-related repetitive TMS reveals distinct, critical roles for right OFA and bilateral posterior STS in judging the sex and trustworthiness of faces. *Journal of Cognitive Neuroscience*, *23*, 2782-2796.

Duranda, K., Gallayb, M., Seigneurica, F., Robichonc, F., & Badouin, J-Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of Experimental Child Psychology, 97*, 14-27

Eger, E., Jedynak, A., Iwaki, T., & Skrandies, W. (2003). Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. *Neuropsychologia, 41,* 808-1

Ekman, P. (1997). Expression or communication about emotion. Uniting psychology and biology: integrative perspectives on human development. (N. Segal, G. E. Weisfeld, & C. C. Weisfeld, eds.). Washington, DC: APA

Ekman, P., & Friesen, W. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, California: Consulting Psychologists Press

Enns, J. T., & DiLollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345-352.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking *Vision Research, 44*, 1321-1331.

Esteves, F., & Öhman, A. (1993). Masking the face: Recognition of emotional facial expressions as a function of the parameters of backward masking. *Scandinavian Journal of Psychology, 34,* 1-18. doi: 10.1111/j.1467-9450.1993.tb01096.x

Harmon, L., & Julesz, B. (1973). Masking in visual recognition: effects of two-dimensional filtered noise. *Science*, *180*, 1194-1197.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223-233

Johnson, M., & Morton, J. (1991). *Biology and cognitive development: the case of face recognition*. Blackwell, Oxford.

Kolers, P. A. (1968): Some psychological aspects of pattern recognition. En P. A. Kolers & M. Eden (Eds.): *Recognizing Patterns*. Boston: MIT Press.

Le Gal, P. M., & Bruce, V. (2002). Evaluating the independence of sex and expression in judgments of faces. *Perception & Psychophysics*, *64*, 230-243.

Lehky, S. R. (2000). Fine discrimination of faces can be performed rapidly. *Journal of Cognitive Neuroscience*, *12*, 848–855.

Liddell, B. J., Williams, L. M., Rathjen, J., Shevrin, H., & Gordon, E. (2004). A temporal dissociation of subliminal versus supraliminal fear perception: An event-related potential study. *Journal of Cognitive Neuroscience*, *16*, 479-486.

Loffler, G., Gordon, G. E., Wilkinson, F., Goren, D., & Wilson, H. R. (2005). Configural masking of faces: evidence for high-level interactions in face perception. *Vision Research, 45*, 2287-2297.

Lundqvist, D., Flykt, A., & Öhman, A (1998). *The Karolinska directed emotional faces—KDEF* [CD-ROM]. Stockholm: Karolinska Institutet, Department of Clinical Neuroscience, Psychology section.

Maxwell, J., & Davidson, R. (2004). Unequally masked: Indexing differences in the perceptual salience of "unseen" facial expressions. *Cognition & Emotion, 18*, 1009-1026.

Milders, M., Sahraie, A., & Logan, S. (2008). Minimum presentation time for masked facial expression discrimination. *Cognition & Emotion, 22*, 63-82.

Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E., & Pernier, J. (2000). Neurophysiological correlates of face gender processing in humans. *European Journal of Neuroscience*, *12*, 303-310.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *Journal of personality and social psychology, 64*, 723-739.

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In M. Lewis and J.M. Haviland (Eds.), *Handbook of emotions* (pp. 511- 536). New York: Guilford Press.

Öhman, A. & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review, 108*, 483-522.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of personality and social psychology*, *80*, 381-396.

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*, 75-92

Pegna, A. J., Khateb, A., Michel, C. M., & Landis, T. (2004). Visual recognition of faces, objects, and words using degraded stimuli: Where and when it occurs. *Human Brain Mapping*, *22,* 300–311.

Pessoa, L., Jappee, S., & Ungerleider, L. G. (2005). Visual awareness and the detection of fearful faces. *Emotion, 5*, 243-247

Pizzagalli, D. A., Regard, M., & Lehmann, D. (1999). Rapid emotional face processing in the human right and left brain hemispheres: An ERP study. *Neuropsychological Report*, *10*, 2691–2698.

Roesch, E. B., Sander, D., Mumenthaler, C., Kerzel, D., & Scherer, K. R. (2010). Psychophysics of emotion: The QUEST for emotion perception. *Journal of Vision, 10*, 1-9. http://journalofvision.org/10/3/4/, doi:10.1167/10.3.4

Schweinberger, S. R., & Soukup, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology: Human perception and performance*, *24*, 1748-1765.

Schyns, P., & Oliva, A. (1999). Dr. angry and Mr. smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition 69*, 243–265.

Sierra-Vázquez, V., & Serrano-Pedraza, I. (2010). Application of Riesz transforms to the isotropic AM-PM decomposition of geometrical-optical illusion images. *Journal of the Optical Society of America A, 27*, 781-796. doi: 10.1364/JOSAA.21000781

Szczepanowski, R., & Pessoa, L. (2007). Fear perception: Can objective and subjective awareness measures be dissociated? *Journal of Vision, 7*(4):10, 1–17.

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, *45*, 174-194.

Whalen, P., Rauch, S., Etcoff, N., McInerney, S., Lee, M. & Jenike, M. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, *18*, 411-418.

(Manuscript received: 26 June 2013; accepted: 4 October 2013)