Routledge
Taylor & Francis Group

Anger superiority effect: The importance of dynamic emotional facial expressions

Francesco Ceccarini[1] and Corrado Caudek[1,2]

[1]Department of Neuroscience, Psychology, Pharmacology, and Child's Health, University of Florence, Firenze, Italy
[2]Centre for Neuroscience and Cognitive Systems@UniTn, Italian Institute of Technology, Rovereto, Italy

A rapid response to a threatening face in a crowd is important to successfully interact in social environments. Visual search tasks have been employed to determine whether there is a processing advantage for detecting an angry face in a crowd, compared to a happy face. The empirical findings supporting the "anger superiority effect" (ASE), however, have been criticized on the basis of possible low-level visual confounds and because of the limited ecological validity of the stimuli. Moreover, a "happiness superiority effect" is usually found with more realistic stimuli. In the present study, we tested the ASE by using dynamic (and static) images of realistic human faces, with validated emotional expressions having similar intensities, after controlling the bottom-up visual saliency and the amount of image motion. In five experiments, we found strong evidence for an ASE when using dynamic displays of facial expressions, but not when the emotions were expressed by static face images.

Keywords: Anger superiority effect; Dynamic faces; Face-in-the-crowd effect; Face perception; Emotion.

Evolutionary arguments may lead us to expect that threatening faces are detected more efficiently among a crowd of distractor faces than non-threatening faces. Because facial threat provides a particularly potent social

Please address all correspondence to Corrado Caudek, Department of Neuroscience, Psychology, Pharmacology, and Child's Health (NEUROFARBA), University of Florence, Via di San Salvi 12, Complesso di San Salvi, Padiglione 26, 50135 Firenze, Italy. E-mail: corrado.caudek@unifi.it

These findings suggest that the ASE might be mediated by mechanisms that are summoned only in part (or not at all) by stimuli lacking the typical dynamic characteristics of facial emotion in natural settings.

signal of warning of potentially aversive consequences, the processing of facial information that conveys threat might recruit greater cognitive resources than the processing of other facial information. In fact, the finding that angry faces are detected more efficiently than happy faces among a crowd of distractors has been reported by several studies (Hansen & Hansen, 1988; Öhman, Lundqvist, & Esteves, 2001). This finding has been termed the "anger superiority effect" (ASE).

Even though the threat detection hypothesis is appealing on the ground of ecological or evolutionary considerations, the evidence supporting this hypothesis has been challenged on multiple fronts. The recent literature has pointed out that the likelihood of finding an ASE in a visual search experiment strongly depends on the presence of low-level visual confounds, a limited set of facial identities, and the use of schematic facial stimuli (Becker, Anderson, Mortensen, Neufeld, & Neel, 2011). When the face stimulus sample comprises a large set of photographs of realistic faces, or when some of the "perceptual confounds" are controlled, the "happiness superiority effect" (HSE) is more likely to be observed that an ASE (Calvo & Nummenmaa, 2008; Juth, Lundqvist, Karlsson, & Öhman, 2005).

Instead of abandoning the threat detection hypothesis altogether, in the present study we considered the possibility that the ASE might be mediated by mechanisms that are summoned only in part (or not at all) by impoverished stimuli. A very important aspect of emotional facial expression that is missing in many tests of the ASE is dynamic information. The goal of the present study was to test the ASE with images of realistic human faces, with validated emotional expressions, with similar intensities of emotional expressiveness, under both static and dynamic conditions, when the bottom-up stimulus saliency and the amount of movement are controlled. Before describing the present experiments, in the next section we briefly examine the literature on the ASE.

EMPIRICAL RESEARCH ON THE THREAT ADVANTAGE EFFECT

Static and dynamic information

In natural conditions, emotional expression can only be transmitted through nonrigid motions resulting from face deformations. Evidence suggesting the importance of dynamic information for identity recognition (LaBar, Crupain, Voyvodic, & McCarthy, 2003; Lander & Chuang, 2005; Lander, Chuang, & Wickham, 2006; O'Toole et al., 2011; O'Toole, Roark, & Abdi, 2002; Pilz, Bülthoff, & Vuong, 2009; Pilz, Thornton, & Bülthoff, 2006; Thornton & Kourtzi, 2002) and emotion recognition (Fujimura & Suzuki, 2010; Kaulard, Cunningham, Bülthoff, & Wallraven, 2012; Nusseck,

Cunningham, Wallraven, & Bülthoff, 2008; Pelphrey, Morris, McCarthy, & LaBar, 2007; Wehrle, Kaiser, Schmidt, & Scherer, 2000) has been provided by behavioural and neuroimaging studies (see Arsalidou, Morris, & Taylor, 2011, for a review). Despite the importance of facial motion for human social cognition, and despite the evidence of an increased neural response evoked by moving compared to static face stimuli (e.g., Schultz & Pilz, 2009), most of the research on the threat detection hypothesis has made use, with a few exceptions, of static images of faces. Static facial stimuli do not reflect the information available from seeing a face in the real world and, thus, may elicit a form of processing that does not reflect the cognitive mechanisms naturally involved in emotion recognition (Foley, Rippon, Thai, Longe, & Senior, 2012).

Becker, Neel, Srinivasan, Neufeld, and Kumar (2012) asked participants to identify the dynamic expression (happy or angry) of one or four faces. The stimuli were singleton faces or crowds of faces all showing the same emotional expression. They found a lateralization effect: Dynamic expressions of happiness were detected faster in the left visual field and dynamic expressions of anger were detected faster in the right visual field (see also Davidson, 1992; Fusar-Poli et al., 2009). Contrary to the threat detection hypothesis, Becker et al. found a consistent advantage for the detection of happiness at the focus of attention for both unprocessed faces and (low-pass and high-pass) filtered face images.

Horstmann and Ansorge (2009) used Musterle faces (Musterle & Rössler, 1986) presented either dynamically or statically. For both static and dynamic stimuli, they found an advantage for the search of a negative target face in a positive distractor crowd over a positive target face in a negative distractor crowd. Moreover, search was more efficient with dynamic stimuli than with static stimuli. However, the Musterle faces produced a larger amount of movement when they expressed a negative rather than a positive emotional expression. In another two experiments, Horstmann and Ansorge controlled the amount of movement of positive and negative emotional expressions. When the schematic faces were matched in terms of the amount of movement, Horstmann and Ansorge did not replicate the result of a more efficient search for angry faces with dynamic stimuli than with static stimuli.

Low-level visual confounds

The ASE might be due to the effect of low-level visual features, rather than to the effect of the facial expressions per se. For example, by comparing search performance across a set of schematic faces that had been used in the previous literature, Horstmann (2009) found that the variability of the search efficiency was mainly explained by the variability of the low-level

properties of the stimuli rather than by the emotional expressions of the faces. Low-level visual confounds have been found in studies employing schematic faces without eyebrows, schematic faces with downturned eyebrows, and images of realistic faces (Becker, Anderson, Mortensen, Neufeld, & Neel, 2011).

With photographs of realistic human faces used as stimuli, Calvo and Nummenmaa (2008) investigated three visual properties that affect the relative discriminability of target and distractors: visual saliency (i.e., stimulus properties such as luminance, colour, and orientation), facial features (significant components of the faces that might guide differently the search for happy and angry target faces with respect to the distractors), and configural information (structural relationships between facial features that may facilitate detection without expression encoding). They found that the main determinant of search efficiency was the visual saliency of some conspicuous facial features. In the case of happy expressions, for example, the facial features in the mouth region were found to be particularly visually salient. Calvo and Nummenmaa concluded that there is no evidence that angry faces are more efficiently detected in a crowd than happy faces, if low-level visual confounds are controlled (see also Becker, Anderson, et al., 2011; Becker, Horstmann, & Remington, 2011; Becker et al., 2012; Coelho, Cloete, & Wallis, 2010; Horstmann, 2009; Hunt, Cooper, Hungr, & Kingstone, 2007; Mak-Fan, Thompson, & Green, 2011; Mermillod, Vermeulen, Lundqvist, & Niedenthal, 2009; Purcell & Stewart, 2010).

The question has thus been raised as to whether the ASE is only determined by perceptual factors, or whether affective factors also play a role. Even though there is still no definite answer to this question, a number of researchers interpret the empirical evidence as suggesting that, if low-level visual confounds are controlled, there is no evidence that angry faces are more efficiently detected in a crowd than happy faces (e.g., Becker, Anderson, et al., 2011; Becker, Horstmann, & Remington, 2011).

Use of schematic faces

The majority of the demonstrations of the threat advantage effect has used drawings of schematic faces (Calvo, Avero, & Lundqvist, 2006; Eastwood, Smilek, & Merikle, 2001; Fox et al., 2000; Horstmann, 2007; Juth et al., 2005; Lundqvist & Öhman, 2005; Mather & Knight, 2006; Öhman et al., 2001; Schubö, Gendolla, Meinecke, & Abele, 2006; Smilek, Frischen, Reynolds, Gerritsen, & Eastwood, 2007; Tipples, Atkinson, & Young, 2002). One issue with schematic faces is that they might transmit negative emotions, but not necessarily a threat. Therefore, it has been suggested that the negative emotions evoked by schematic faces might not be relevant for

testing the threat detection hypothesis (Becker, Anderson, et al., 2011).[1] Another issue with schematic faces is that there is no perceptual variability in the stimulus displays, which is completely unrealistic.

THE CURRENT STUDY

In Experiments 1A, 1B, 3A, and 3B of the present study, we used a search-asymmetry design (i.e., one angry face in a crowd of happy faces vs. one happy face in a crowd of angry faces) with either dynamic or static facial expressions of emotions. We decided to use a search-asymmetry design because it has been used most extensively in the literature on the face-in-the-crowd effect and because it tests the threat-advantage hypothesis most directly. In the search-asymmetry design with happy and angry faces, in fact, the stimuli that are contrasted constitute the extreme values on the dimension of facial threat (Horstmann, 2007; Horstmann & Bauland, 2006).

Note that in the two most notable studies concerning the ASE with dynamic facial emotional expressions (see Static and dynamic information in the previous section), either participants did not perform a visual search task, or the stimuli were not images of realistic human emotional faces. In Experiment 2, neutral faces were used as a baseline condition. We took the following steps to address some of the potential confounds that have been implicated in the previous literature on the ASE.

External validity

We used validated images of emotional expressions selected from the Radboud Faces Database to generate ecologically valid stimuli (Langner et al., 2010).

Bottom-up visual saliency

With stimuli derived from images of realistic human faces, it is important to control for the effects of low-level visual features that may differ across happy and angry faces (Batty, Cave, & Pauli, 2005; Calvo & Esteves, 2005).

---

[1] Researchers in the field have different opinions on the validity of the schematic faces for testing the ASE. Evidence in favour of the use of schematic faces has been provided, for example, by Öhman et al. (2001). In their fifth experiment, they evaluated whether the search asymmetry favouring angry faces is due to the threat value of the schematic faces, or whether it reflects a more general bias for negative information. An angry, happy, sad, or "scheming" face was used as the target among neutral distractors. In both RTs and error rates, Öhman et al. found an advantage for the angry target face as compared to the other three possible target faces. These results support the hypothesis that the ASE is specifically produced by the threat value, also in the case of schematic faces (but see Appendix B).

We addressed this issue by estimating the bottom-up visual saliency according to the computational model proposed by Itti and Koch (2000, 2001).[2]

Amount of motion

We selected the stimulus displays to match the amount of movement of happy, angry, and neutral faces (see Appendix A).

Natural variability of emotional expressions

Many of previous studies employed a rather small sample of facial identities. When a small number of face identities is presented in each stimulus display, participants can learn a single identical feature that discriminates distractors from targets, especially if the experiment comprises a very large number of trials. Therefore, it is important that distractors and target faces have different identities and that the distractor faces show some variability, so as to represent the natural variability of emotional expressions (Becker, Anderson et al., 2011; Calvo & Nummenmaa, 2008).

In the present study, we addressed this issue by presenting nine facial identities in each stimulus display and by asking participants to perform a moderate number of trials. Our static stimuli are thus similar to those used by Pinkham, Griffin, Baron, Gur, and Sasson (2010), who found an ASE with static photographs of real faces. Our stimuli also reduce the possibility of grouping strategies. In the case of schematic faces, in fact, has been shown that individual distractors can be selected in chunks (not in a strictly serial fashion), in such a way as to enable participants to process and reject in parallel multiple items (Becker, Horstmann, & Remington, 2011). This cannot be done, however, when the distractors are not all identical to each other.

EXPERIMENTS 1A AND 1B

A search-asymmetry design (a target angry face surrounded by happy nontarget faces, or a target happy face surrounded by angry nontarget faces) was used to determine whether dynamic facial expressions might provide a more appropriate stimulus to examine the ASE than static stimuli. The stimulus sets representing anger and happiness were equated in terms of

---

[2] The saliency map can successfully predict a large fraction of fixated locations during inspection of complex scenes (Cerf, Harel, Einhäuser, & Koch, 2008; Foulsham & Underwood, 2008; Kowler, 2011; Mannan, Kennard, & Husain, 2009; Parkhurst, Law, & Niebur, 2002) and it can also predict the visual search results obtained with some of the schematic faces used in the literature (see Appendix B).

bottom-up saliency as measured by the saliency maps produced by the Itti-Koch model (Itti & Koch, 2000, 2001; Koch & Ullman, 1985) and by the amount of image motion. In Experiment 1A, static and dynamic trials were blocked, with the order of blocks counterbalanced across participants. In Experiment 1B, static and dynamic trials were randomly interleaved.

Method

Participants.   A total of 25 undergraduate students from Florence University participated in the experiments, 16 women and nine men, with a mean age of 26.9 years ($SD = 4.1$). Ten undergraduates participated in Experiment 1A and 15 undergraduates participated in Experiment 1B. All participants were naïve to the purpose of the study. All participants had normal or corrected-to-normal vision. Participation was voluntary.

Apparatus.   Stimulus presentation and data acquisition were conducted using a 486-based PC-compatible computer connected to a 17-inch video monitor operating at 60 Hz. A custom Visual C++ program was used for stimulus presentation and response recording.

Stimuli.   Stimuli consisted of three-dimensional (3D) models of faces each measuring 200×200 pixels. The face images were selected from the Radboud Faces Database (Langner et al., 2010), which is a standardized set of face images that display facial expressions based on prototypes from the Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002). Nine different identities were used. Each image was cropped to remove hair and background. In the Radboud database, the happy faces are represented with open mouth smiles, whereas the angry faces are represented with the mouth closed. Images of faces that have an open mouth relative to close-mouthed faces yield a simple visual feature (visible teeth) that can drive efficient search because it easily differentiates the target from the distractors (Becker, Anderson, et al., 2011; Purcell, Stewart, & Skov, 1996).[3] To address this issue, we modified the happy face images as follows. The stimulus generation started with the creation of 54 3D models of faces (happy or angry) representing 27 different face identities. From the Radboud Faces Database, we selected the face images that received similar evaluations in terms of intensity of the expression, clarity of the expression, genuineness of the expression, and valence. The 27 3D models of faces expressing happiness

---

[3] The influence of teeth visibility on efficiency of visual search has been recently studied by Horstmann, Lipp, and Becker (2012). They found that (1) the effect of teeth visibility on search efficiency was as strong as the effect of target presence, and (2) the effect of teeth visibility on search efficiency was stronger than the effect of facial emotion.
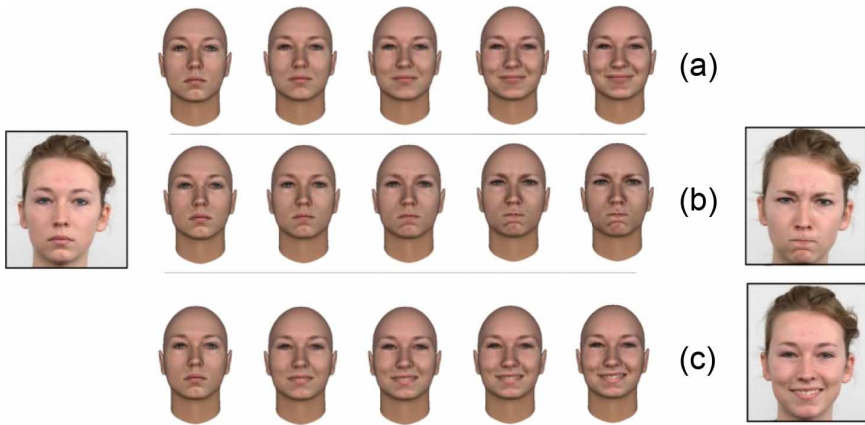
Figure 1. Example of stimulus generation procedure. The three faces enclosed in the rectangular frames have been selected from the Radboud Faces Database and show the same identity with three different emotional expressions: neutral, happy, and angry. These images were transformed to remove hair and then morphed to obtain a smooth transition between the neutral expression and the full-emotion expression. The figure shows some representative frames of the morph continua in the case of (a) a transition between a face with a neutral expression and a face with an emotional expression of happiness with the mouth closed (Experiments 1A, 1B, and 2), (b) a transition between a face with a neutral expression and a face with an emotional expression of anger with the mouth closed (Experiments 1A, 1B, and 2), and (c) a transition between a face with a neutral expression and a face with an emotional expression of happiness with the mouth open (Experiments 3A and 3B). To view this figure in colour, please see the online issue of the Journal.

with an open mouth were then transformed with the 3d Studio Max software in order to create 3D face models expressing happiness with a closed mouth (see Figure 1a).

In the dynamic condition, neutral faces gradually expressed positive (happiness) or negative (anger) emotions. The frame sequence was created by morphing between a face with a neutral expression and a face with an emotional expression of anger or happiness. Pairs of neutral and happy, or neutral and angry faces were morphed into a continuum of 16 steps. Figures 1a and b show the states of the transition between the neutral (left) and the negative or positive (right) emotional expressions. Corresponding static stimuli were created by repeating the final frames of the video sequences.

The total duration of the frame sequence was 2000 ms. Within this temporal window, a neutral face was displayed for 300 ms, followed by the morphing transition between the neutral face and the final expressive face (each of the 16 faces of a morph continuum was shown in succession and remained on the screen for 1/30 s for a total of 533 ms), and by the final expression of the face (1167 ms). This procedure, with components that change in a linear fashion, has the advantage of allowing a precise control of

the timing of the change, without sacrificing the realism of the expressive dynamics (e.g., Becker et al., 2012). The duration of the temporal unfolding of facial expressions of emotion (533 ms) was in line with other experiments generating dynamic facial expressions with methods similar to the present study (e.g., Arsalidou et al., 2011: 300 ms; Becker et al., 2012: 105 ms; Horstmann & Ansorge, 2009: 500 ms; Schultz & Pilz, 2009: 1040 ms). In the static condition, the full angry or happy emotional expressions were presented from the beginning of the trial.

All displays measured 1280×1024 pixels and were centred in the middle of the monitor without a visible border. The typical viewing distance was 60 cm.

Intensity of emotional expressiveness: FaceReader.   The facial expressions of the selected 27 face identities were coded by means of the Noldus FaceReader software (Noldus Information Technologies, 2012). FaceReader uses the locations of 55 key points and the texture of the face to measure the intensity of the emotional expression, as compared with the facial emotional prototypes of Ekman and Friesen (1975). FaceReader has shown to have high convergent validity with the ratings made by human FACS experts (Chentsova-Dutton & Tsai, 2010; den Uyl & van Kuilenburg, 2005; Truong, van Leeuwen, & Neerincx, 2007). A score of emotional expressiveness ranging from 0 to 1 was assigned to each of the 27 face identities. We discarded the nine face identities that obtained an emotional expressiveness score lower than .95.

The 18 3D models selected in this manner were then processed with the 3dStudio Max software in order to equate lighting intensity and direction. These 3D face models were then exported in PNG format and successively transformed in movie sequences (30 fps) by means of the Flash CS5 software.

Amount of image motion.   We measured the amount of image motion to insure that it was approximately the same with negative and positive emotional expressions (see Horstmann & Ansorge, 2009). In our stimuli, the face images could differ only in terms of their expression: The face orientation, the ambient illumination, and the face dimensions were kept constant. The amount of movement was evaluated by comparing the first (neutral) and the last (full emotion) frame of the video sequence for each face identity and each expression, as indicated in Appendix A.

For each of 18 face identities selected as indicated earlier, we measured the amount of image motion generated by the temporal unfolding of a happy or angry emotion. We then selected the nine face identities that produced equivalent amounts of image motion (see Equation A3) when expressing the two emotions, $t(15.97) = 0.008$, $p = .99$. A sample stimulus display comprising the nine face identities used in Experiments 1A and 1B is shown in Figure 2.
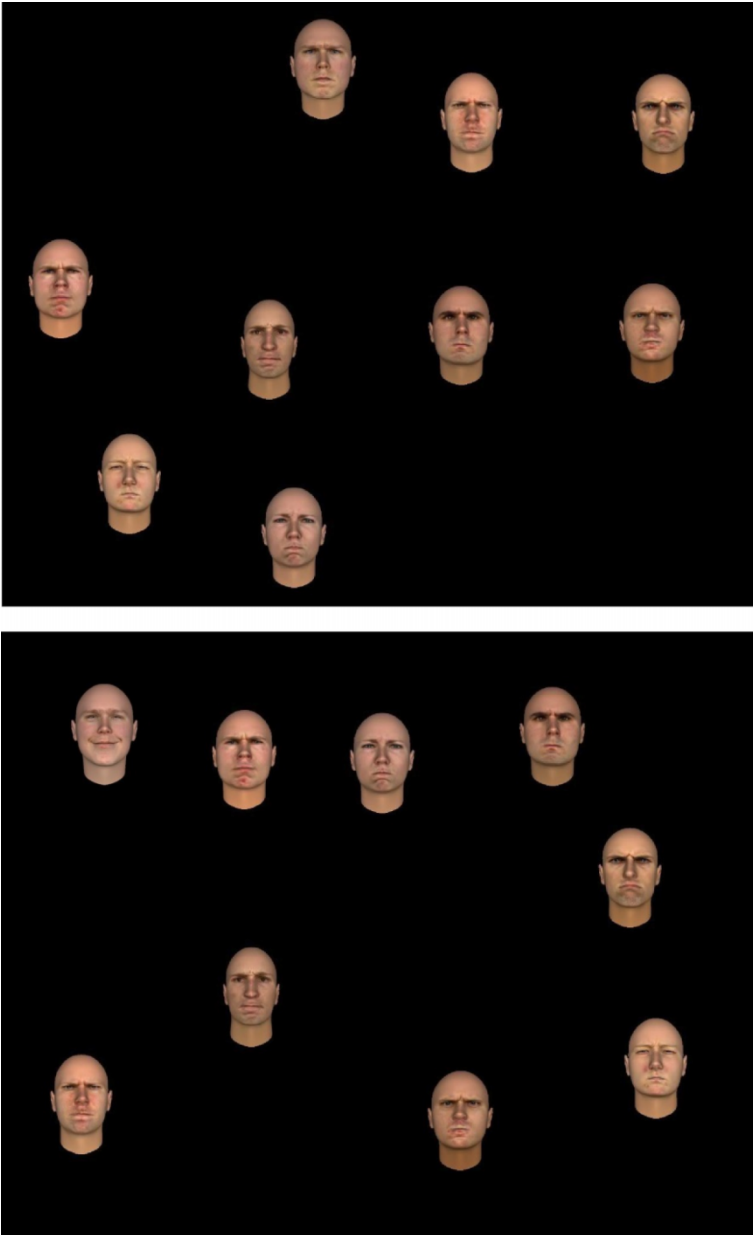
Figure 2. Examples of stimulus displays. Top: target-absent display (all-angry crowd). Bottom: target-present display (a happy face in an angry crowd). To view this figure in colour, please see the online issue of the Journal.

Intensity of emotional expressiveness: Human judges.   A further check was carried out by asking human observers to judge the intensity of the emotional facial expressions of the eighteen selected images (nine face identities). A total of 107 undergraduate students participated in this evaluation. Static images of the full positive or negative emotional facial expressions were successively presented for 15 s in random order. Participants provided a judgement of the emotional intensity of each face on a 4-point scale (1 = low, 2 = medium, 3 = high, 4 = very high). The results indicate that judgements of emotional intensity did not significantly vary across emotional valence (happy: mean = 3.05, SE = 0.015; angry: mean = 3.05, SE = 0.018), $z = 0.13$, $p = .90$.

Bottom-up visual saliency.   We measured bottom-up visual saliency by means of a well-used model of visual attention (Itti & Koch, 2000, 2001; Koch & Ullman, 1985). The model assumes that the allocation of visual attention is driven by stimulus saliency in a bottom-up fashion and analyses natural images by extracting low-level features such as intensity, colour, and orientation at a range of spatial scales. To optimize the detection of local feature differences, these features are converted to centre-surround representations. From these representations, separate "conspicuity" maps are created. The conspicuity maps are then combined to form one saliency map that guides the attention focus.

The 18 selected face images were used to generate 1600 displays comprising one target face embedded in a crowd of seven face distractors having a different emotional expression (i.e., one angry face and seven happy faces, or one happy face and seven angry faces). Each display was divided into a 2 × 4 matrix of regions of 512 × 320 pixels each (Figure 3). Each of the nine happy faces and of the nine angry faces employed in Experiments 1A and 1B was used as target in eight different displays (once for each of the eight cells of the matrix). In each display, the seven distractor faces were randomly chosen among the nine faces with opposite emotional valence and were randomly positioned in the remaining cells of the matrix. These images were then processed with the SaliencyToolbox 2.2 for MAT-LAB (Walther & Koch, 2006; http://www.saliencytoolbox.net/). For the purpose of this study, the standard settings were used. For each display, a saliency index was computed by measuring the total activation within the region in which the target face was located (Calvo & Nummenmaa, 2008; Humphrey, Underwood, & Lambert, 2012).[4] This

---

[4] Note that the saliency score computed in this manner indicates how different the target face is from the distractors, not how much the target stands out from the background. In other words, the algorithm produces different saliency scores if a constant target face is surrounded by distractors or if it is shown in isolation in the same location.
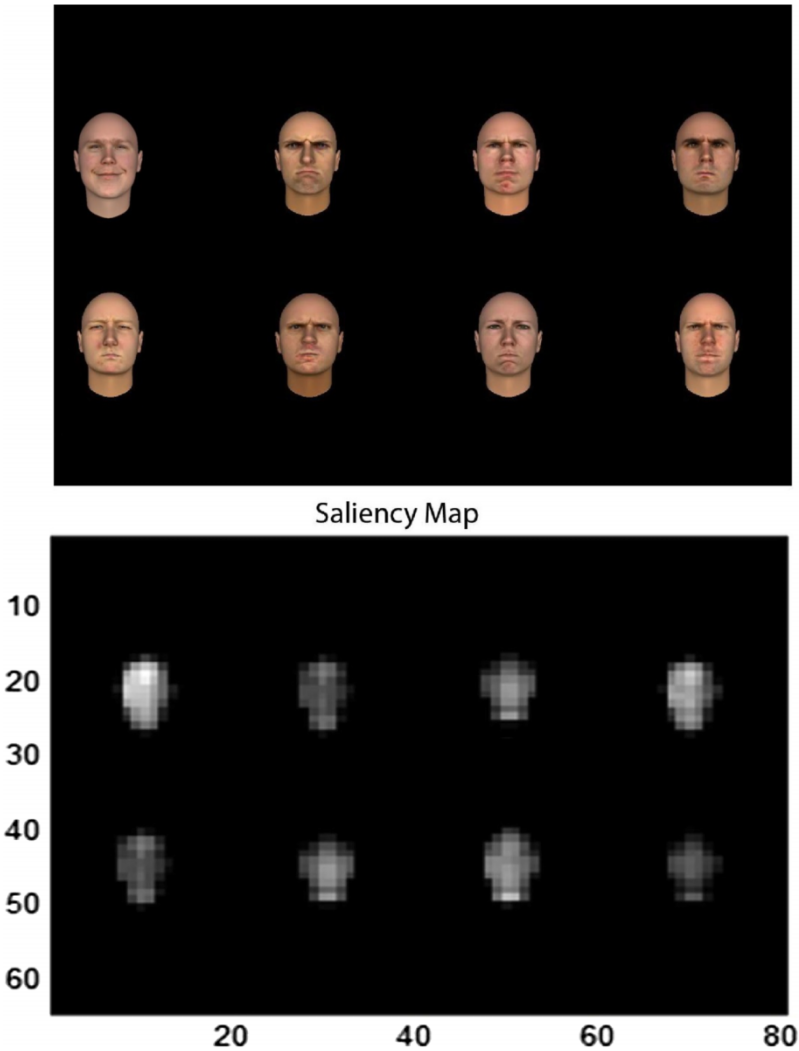
Figure 3. Saliency map. Example of input image (top) used to compute the saliency map (bottom) according to the algorithm of Itti and Koch (2000, 2001). To view this figure in colour, please see the online issue of the Journal.

analysis showed that the faces used in Experiments 1A and 1B did not differ in terms of the bottom-up saliency when displaying a happy or an angry expression, $t(1591) = 0.35$, $p = .73$.[5]

[5] To verify that this analysis was performed correctly, we replicated the results reported by Calvo and Nummenmaa (2008)—see Appendix B.

Design. Stimulus (static vs. dynamic) and crowd (happy crowd vs. angry crowd) were within-subject variables. The dependent variable was the detection latency.

Each session consisted of 216 trials presented in a random order. One-quarter (54) of the trials were target-present trials composed of eight faces showing a happy expression and one face showing an angry expression, one-quarter of the trials were target-present trials composed of eight faces showing an angry expression and one face showing a happy expression, one-quarter of the trials were target-absent trials composed of nine faces showing a happy expression, and one-quarter of the trials were target-absent trials composed of nine faces showing an angry expression. In half of the trials, facial expressions were dynamic; in the other half of the trials, facial expressions were static. In the target-present trials of each experimental session, every face identity was used as target three times. The positions of the face images in the stimulus displays were determined randomly on each trial with the constraint that the minimum distance between the (200×200 pixels) regions containing each face was at least 10 pixels.

Experiment 1A comprised two blocks, one with dynamic faces, the other with static faces (see Horstmann & Ansorge, 2009). The two blocks were separated by a break of at least 30 minutes. The order of block presentation was counterbalanced across participants. The two blocks of trials were preceded by 20 practice trials. Stimulus exposure was terminated by the participant's response.

In Experiment 1B, trials with static and dynamic displays were randomly interleaved into two blocks separated by a break of at least 30 minutes. The two blocks of trials were preceded by 20 practice trials. The purpose of Experiment 1B was to minimize the likelihood of participants using different response strategies for static and dynamic displays.[6] To minimize local processing strategies, stimulus exposure was limited to 2000 ms (see Caudek & Monni, 2013).

Procedure.   Each trial began with the presentation of a fixation cross (500 ms) followed by the presentation of a random arrangement of nine expressive faces. Participants were asked to indicate with a keypress whether all faces showed the same expression or whether one face showed an expression differing from the others. Participants were instructed to perform the task as quickly and accurately as possible. No feedback was provided for correct or incorrect responses.

---

[6] It is important to keep participants' expectations constant across conditions because visual search is a highly contextualized process and search performance depends strongly on the specific demands of the task (e.g., Frischen, Eastwood, & Smilek, 2008).

Statistical analyses.   To analyse the reaction times, we fitted fully parametrized (e.g., Kliegl, Wei, Dambacher, Yan, & Zhou, 2011) linear mixed effects (LME) models to the data (Pinheiro & Bates 2000) by using the lme4 package in the R statistical programming environment (Bates, Maechler, & Dai, 2009). LME models offer several advantages over ordinary least squares regression for the analysis of reaction-time data, in particular: increased statistical power, a more flexible method of dealing with missing data, and a better modelling of heteroscedasticity (Baayen, Davidson, & Bates, 2008). Importantly, by treating the stimuli as a random factor in the analysis, LME models allow a more robust generalization of the results across stimuli (Caudek & Domini, 2013; Judd, Westfall, & Kenny, 2012). Target-present and target-absent trials were analysed separately with crowd valence (happy, angry) and stimulus (static, dynamic) as fixed effects, with subjects and target images as crossed random effects (Baayen, 2008).

Following the procedure indicated Baayen and Milin (2010), outliers were removed by a two-step procedure. First, extremely short RTs (button presses within 200 ms of stimulus onset) and extremely long latencies (exceeding 6 s) were excluded. This screening produced a loss of less than 1% of the data. Second, we computed the standardized residuals for the full LME model. All statistical tests were then performed by refitting the model after removing data points with absolute standardized residuals exceeding 2.5 standard deviations. This produced a loss of 3.4% of the data in target-present trials and 2.9% of the data in target-absent trials. Evidence that normality was not violated comes from the observation that residuals were approximately normally distributed. We found no signs of heteroscedasticity.

Analysis of errors on the search task was performed with generalized linear mixed-effects models with binomial error structure after coding error and correct responses with 1 and 0, respectively. The GLME models had the same fixed-effects used in the analysis of the RTs, with the inclusion of subject and target images as crossed random effects.

All analyses were performed by identifying the maximal random-effects structure justified by the data, so as to obtain the best protection against the Type I error (see Barr, Levy, Scheepers, & Tily, 2013). This was achieved by using LME models including by-subject and by-item random intercepts, by-subject random slopes for all within-subjects factors, and (when this was possible by LME model convergence) by-subject random slopes for the interactions of the within-subjects factors. We evaluated significance by computing the deviance statistic (minus 2 times the log-likelihood). Change in deviance is distributed as chi-square, with degrees of freedom equal to the number of parameters deleted from the model (see Baayen et al., 2008).

Results—Experiment 1A

Response latencies.   Mean detection latency on "correct detection" trials as a function of crowd, stimulus, and target presence is shown in Figure 4. On average, participants responded 547 ms faster in target-present than in target-absent trials, $\chi_1^2 = 37.5$, $p = .001$.

Target-present trials.   The Stimulus $\times$ Crowd interaction was statistically significant, $\chi_1^2 = 6.52$, $p = .011$. Under dynamic conditions, average RTs were 257 ms faster for angry-face targets among happy-face distractors than the opposite (happy-face targets among angry-face distractors), $\chi_1^2 = 10.3$, $p = .002$. Under static conditions, the search for an angry face within a crowd of happy faces did not produce RTs significantly different from those of the search for a happy face within a crowd of angry faces, $\chi_1^2 = 0.19$, $p = .66$.

Target-absent trials.   The effect of the variable crowd was statistically significant, $\chi_1^2 = 15.6$, $p = .001$: Average RTs were 263 ms longer for
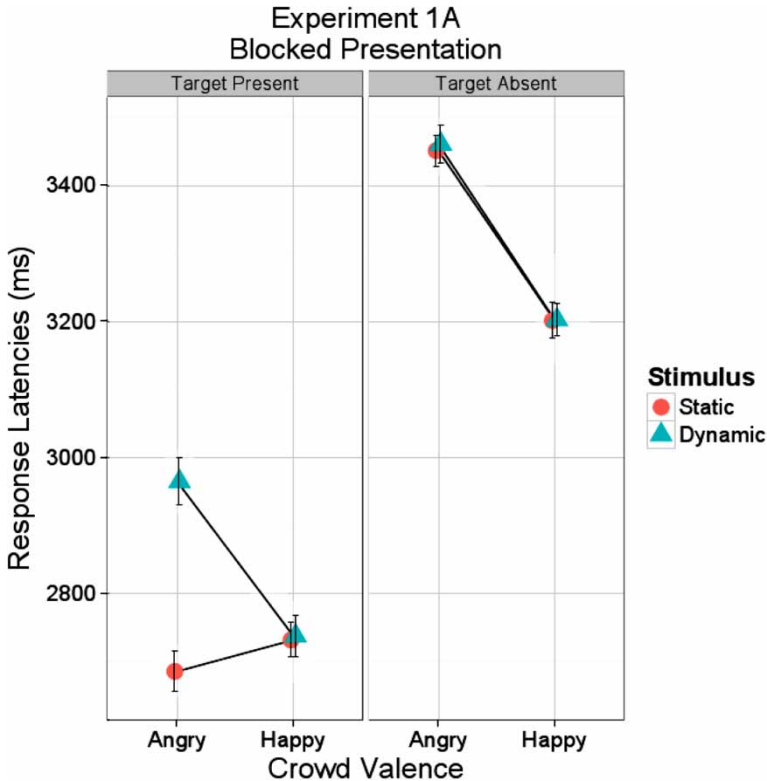

Figure 4.   Experiment 1A. Response latencies as a function of target (present, absent), crowd valence (angry, happy), and stimulus (circle: static; triangle: dynamic). Error bars represent ±one standard error. To view this figure in colour, please see the online issue of the Journal.

TABLE 1
Mean percentage errors for target-present and target-absent trials in the visual search task of Experiment 1A as a function of stimulus (static, dynamic) and crowd valence (angry, happy)

Target	Stimulus	Crowd valence	% errors
Present	Static	Angry	0.8
		Happy	0.4
	Dynamic	Angry	0.4
		Happy	0.4
Absent	Static	Angry	0.0
		Happy	0.0
	Dynamic	Angry	0.4
		Happy	0.0

all-angry crowds than for all-happy crowds. Neither the interaction Stimulus × Crowd nor the effect of the variable stimulus were statistically significant.

Error analysis.   Mean percentage errors are presented in Table 1 as a function of target (present, absent), stimulus (static, dynamic), and crowd valence (happy, angry).

Target-present trials.   The mean error rate for target-present trials was less than 1%. The error rates were neither affected by stimulus nor by crowd nor by their interaction, $\chi_3^2 = 0.55$, $p = .91$.

Target-absent trials.   The mean error rate for target-absent trials was less than 1%. The error rates were neither affected by stimulus nor by crowd nor by their interaction, $\chi_3^2 = 2.78$, $p = .43$.

Results—Experiment 1B

Response latencies.   Mean detection latency on "correct detection" trials as a function of crowd valence, stimulus, and target presence is shown in Figure 5. On average, participants responded 616 ms faster in target-present than in target-absent trials, $\chi_1^2 = 27.2$, $p = .001$.

Target-present trials.   The Stimulus × Crowd interaction was statistically significant, $\chi_1^2 = 4.36$, $p = .037$. Under dynamic conditions, average RTs were 203 ms faster for angry-face targets among happy-face distractors than the opposite (happy-face targets among angry-face distractors), $\chi_1^2 = 13.6$, $p = .001$. Under static conditions, the search for an angry face within a crowd of happy faces did not produce RTs significantly different from those of the search for a happy face within a crowd of angry faces, $\chi_1^2 = 1.4$, $p = .24$.

Target-absent trials.   The effect of the variable crowd was statistically significant, $\chi_1^2 = 26.1$, $p = .001$: Average RTs were 182 ms longer for all-angry crowds than for all-happy crowds. Neither the interaction
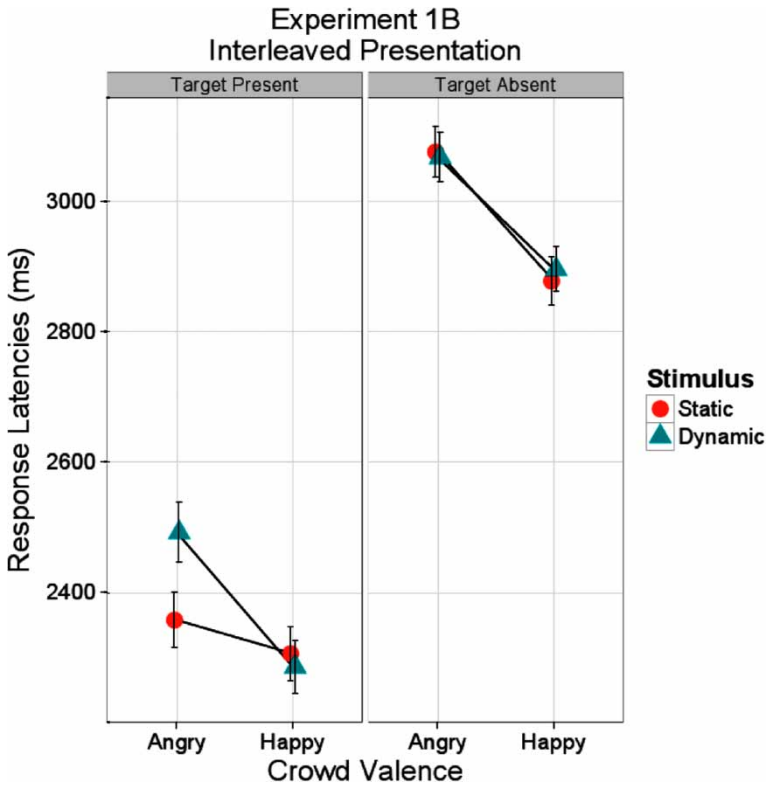
Figure 5.   Experiment 1B. Response latencies as a function of target (present, absent), crowd valence (angry, happy), and stimulus (circle: static; triangle: dynamic). Error bars represent ±one standard error. To view this figure in colour, please see the online issue of the Journal.

Stimulus × Crowd nor the effect of the variable stimulus were statistically significant.

Error analysis.   Mean percentage errors are presented in Table 2 as a function of target (present, absent), stimulus (static, dynamic), and crowd valence (happy, angry).

Target-present trials.   The mean error rate for target-present trials was 14.4%. The error rates was neither affected by stimulus nor by crowd nor by their interaction, $\chi_7^2 = 1.28$, $p = .99$.

Target-absent trials.   The mean error rate for target-absent trials was 3.1%. The interaction Crowd × Stimulus was not statistically significant, $\chi_1^2 = 2.17$, $p = .14$. The effect of the variable stimulus was not statistically significant, $\chi_1^2 = 0.42$, $p = .52$. The effect of the variable crowd was statistically

TABLE 2
Mean percentage errors for target-present and target-absent trials in the visual search task of Experiment 1B as a function of stimulus (static, dynamic) and crowd valence (angry, happy)

Target	Stimulus	Crowd valence	% errors
Present	Static	Angry	11.7
		Happy	16.3
	Dynamic	Angry	14.1
		Happy	15.7
Absent	Static	Angry	5.0
		Happy	1.6
	Dynamic	Angry	5.4
		Happy	0.5

significant, $\chi_1^2 = 13.9$, $p = .001$. The error rates were 5.2% and 1.1% for angry and happy crowds, respectively.

Comparison of Experiments 1A and 1B

Response latencies. The mean RT for correct trials was longer in Experiment 1A (blocked presentation) than in Experiment 1B (interleaved presentation), both on target-present trials (2765 ms vs. 2414 ms), $\chi_1^2 = 4.4$, $p = .036$, and target-absent trials (3325 ms vs. 2972 ms), $p = .020$, $\chi_1^2 = 5.39$

Error analysis. The error rates were overall smaller in Experiment 1A than in Experiment 1B, both on target-present trials (0.5% vs. 14.3%), $\chi_1^2 = 17.1$, $p = .001$, and target-absent trials (0.0% vs. 2.8%), $p = .001$, $\chi_1^2 = 22.8$

Discussion

In Experiments 1A and 1B we found evidence of an ASE, but only in the case of dynamic faces. In the target-present trials, the mean search latencies were significantly longer for a dynamic happy face within a crowd of angry faces than for a dynamic angry face within a crowd of happy faces. For static facial expressions, instead, the search for a happy or angry target face produced similar response latencies.

In both target-present and target-absent trials, mean response latencies were longer in Experiment 1A (blocked presentation) than in Experiment 1B (interleaved presentation). This result can be attributed to the different stimulus exposure times in the two experiments: In Experiment 1A, stimulus presentation time was unlimited; in Experiment 1B, instead, the stimulus was

presented for a maximum of 2000 ms. Consistently, the number of errors was smaller in Experiment 1A than in Experiment 1B.

In Experiment 1A, error rates were for the most part negligible. In the target-present trials of Experiment 1B, the error rates were not modulated by crowd valence (happy, angry) nor by the stimulus type (static, dynamic). In target-absent trials, accuracy was lower in all-angry trials as compared to all-happy trials (see also Pinkham et al., 2010).

In a search-asymmetry design is important to consider the target-absent trials, because the effects of attentional guidance by the target is confounded by the speed of the serial scanning through the crowd of distractors. In both dynamic and static trials, we found that all-angry crowds were searched more slowly than all-happy crowds. This result is consistent with previous findings, but it is problematic in a search-asymmetry design, because it might suggest that the search results depend not only on the processing of the target, but also on the differences in dwell time associated with the distractors (e.g., Frischen et al., 2008; Hansen & Hansen, 1988; White, 1995). For example, Fox Russo, and Dutton (2002) interpreted their evidence in favour of the "threat hypothesis" as due to the enhanced detection of the angry target face as well as the longer dwell time for angry faces relative to happy or neutral faces. In the present case, the "delayed disengagement" hypothesis can explain our results only in part, because we found evidence of an ASE only in the dynamic trials, even though the dwell times on distractors were similar in dynamic and static target-absent trials. Nevertheless, a more direct test of the "delayed disengagement" hypothesis requires an experimental design that examines the effects of happy, angry, and *neutral* distractors on the visual search for an emotional (happy, angry) or *neutral* target singleton. This was the purpose of Experiment 2.

EXPERIMENT 2

Experiment 2 introduced a neutral baseline in order to pursue two goals. First, we wished to determine whether the results of Experiments 1A and 1B might be accounted for by an enhanced attentional dwell time on threatening faces, rather than by a facilitated orienting towards threatening stimuli (Fox et al., 2002). If participants take longer to disengage from angry faces, then response times should be longer when searching for a neutral singleton among angry distractors than when searching for an angry singleton among neural distractors. Second, we intended to compare search efficiency between happy and angry singleton targets within crowds of neutral faces.
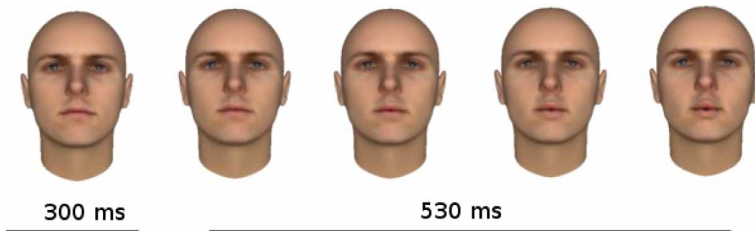
Figure 6. A dynamic face with neutral expression (Experiment 2). The figure shows some representative frames of the transitions between the face shapes associated with the spoken phoneme /W/. To view this figure in colour, please see the online issue of the Journal.

Method

Participants.   A total of 11 undergraduate students from Florence University participated in the experiment, seven women and four men, with a mean age of 23.4 years ($SD = 2.1$). All participants were naïve to the purpose of the study. All participants had normal or corrected-to-normal vision. Participation was voluntary. None of them had participated in Experiments 1A and 1B.

Apparatus.   The apparatus was the same as in Experiment 1A.

Stimuli.   The function PhotoFit SDK of the FaceGen software was used to generate realistic images of the spoken phoneme /W/ for the nine face identities of Experiment 1A (see Figure 6). The frame sequences representing dynamic neutral faces were generated by a linear morphing between (1) the untransformed images of the neutral faces selected from the Radboud database, and (2) the images that had been transformed as indicated earlier. Otherwise, the stimuli were the same as in Experiment 1A.

Intensity of emotional expressiveness: FaceReader.   An analysis performed with the Noldus FaceReader software indicated that the articulation of the spoken phoneme /W/ did not alter the emotional expression of the neutral faces, $t(16) = -0.0123$, $p = .99$.

Amount of image motion.   By following the same procedure as in Experiment 1A, we measured the amount of image motion generated by the 27 face images (9 face identities $\times 3$ face expressions) used in Experiment 2—see Equation A3. We found no evidence that the amount of image motion differed across the three face expressions (happy, angry, and neutral), $F(2, 24) = 0.001$, $p = .99$.

Bottom-up visual saliency.   With the 27 face images used in Experiment 2, we generated 2400 displays comprising a target face (happy, angry, or neutral) embedded in a crowd of seven face distractors having a different

emotional expression (happy, angry, or neutral). These images were then processed with the SaliencyToolbox 2.2 for MATLAB. By following the same procedure as in Experiments 1A and 1B, we found that the faces used in Experiment 2 did not differ on average in terms of bottom-up saliency, when displaying a happy, angry, or neural expression, $F(2, 2376) = 0.22$, $p = .80$.

Design.   The design of Experiment 2 replicated the dynamic condition of Experiment 1A, except that the face stimuli displayed angry, happy, or neutral expressions. Condition (8h1a: eight happy distractor faces and one angry target face; 8a1h: eight angry distractor faces and one happy target face; 8n1a: eight neutral distractor faces and one angry target face; 8a1n: eight angry distractor faces and one neutral target face; 8n1h: eight neutral distractor faces and one happy target face; 8h1n: eight happy distractor faces and one neutral target face; 9a: nine angry distractor faces; 9h: nine happy distractor faces; 9n: nine neutral distractor faces) was the only within-subject variable. Each session consisted of 324 trials presented in a random order (36 trials for each of the nine levels of the variable condition).

Procedure.   The procedure was the same as in Experiment 1A.

Results

Response latencies.   On "correct detection" trials, participants were, on average, 1107 ms faster in target-present than in target-absent trials, $\chi^2_1 = 28.1$, $p = .001$.
Target-present trials.   The effect of condition was statistically significant, $\chi^2_5 = 16.4$, $p = .006$. Table 3 reports the planned comparisons that were performed with the R add-on package multcomp to control the overall Type I error rate (Hothorn, Bretz, & Westfall, 2008).
From the obtained results, we can highlight the following points (see Figure 7). (1) The 8n1a–8n1h comparison indicates that angry target faces

TABLE 3
Planned comparisons for Experiment 2

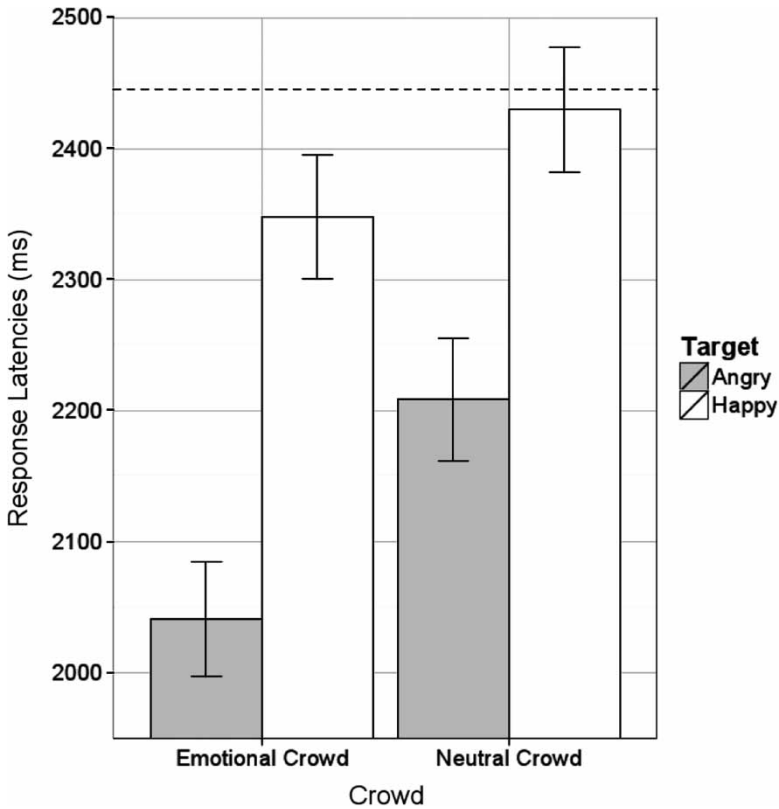| Comparison	Estimate	SE	z-value	p-value
8a1h–8a1n = 0	131.4	95.4	1.38	.563
8h1a–8h1n = 0	−361.2	119.4	−3.02	.012
8n1a–8n1h = 0	−229.5	83.5	−2.75	.030
8a1h–8h1a = 0	327.4	68.1	4.81	.001
8a1n–8n1a = 0	281.7	99.2	2.84	.023
8h1n–8n1h = 0	−45.4	66.3	−0.68	.953

Figure 7. Experiment 2. Response latencies on "correct detection" trials as a function of crowd (emotional, neutral) and target (angry, happy). The dashed line represents the mean latencies in the 8a1n and 8h1n conditions (8a1n: 2481 ms, $SE = 49$; 8h1n: 2359 ms, $SE = 43$). Error bars represent $\pm$ one standard error.

were detected 229 ms faster than happy faces when they were embedded in a neutral crowd. (2) For a crowd of angry faces, the happy target was not processed significantly faster than a neutral target (8a1h–8a1n). Instead, for a crowd of happy faces, an angry target was processed significantly faster than a neutral target (8h1a–8h1n). (3) The 8a1n–8n1a comparison reveals an advantage for the angry target faces, whereas the corresponding asymmetry (8h1n–8n1h) did not show an advantage for the happy target faces. The 8a1h–8h1a comparison replicates the advantage for angry over happy target faces in an emotional crowd that was found in Experiments 1A and 1B. (4) In a further analysis, we examined the trials with an emotional (8h1a, 8a1h) or neutral (the average of the 8a1n and 8h1n conditions) singleton target within an emotional crowd. We found an advantage for the angry target face

TABLE 4
Mean percentage errors for target-present trials in the visual search task of Experiment 2 as a function of crowd valence (happy, neutral, angry) and target valence (happy, neutral, angry)

Crowd valence	Target valence	% errors
Happy	Angry	8.8
Happy	Neutral	12.5
Neutral	Angry	5.0
Neutral	Happy	7.7
Angry	Happy	16.3
Angry	Neutral	11.6

(8h1a, average $= 2040$ ms, $SE = 40$) over the neutral target face (average of the 8a1n and the 8h1n conditions $= 2433$ ms, $SE = 33$), $t = 4.49$, $p = .001$, but no statistically significant difference between the happy target face (8a1h, average $= 2368$ ms, $SE = 45$) and the neutral target face, $t = 0.89$, $p = .38$.[7] *Target-absent trials.* On correct trials, the mean RT values were 3250 ms ($SE = 35$), 3482 ms ($SE = 39$), and 3435 ms ($SE = 41$) for happy, angry, and neutral crowds, respectively. Three planned comparisons indicated that happy crowds were searched 259 ms faster than angry crowds, $z = 4.28$, $p = .001$. Neither the angry versus neutral, $z = 0.51$, $p = .863$, nor the happy versus neutral, $z = -2.09$, $p = .087$, comparisons reached statistical significance.

Error analysis. Mean percentage errors are presented in Table 4 as a function of target valence (happy, angry, neutral), and crowd valence (happy, angry, neutral).
Target-present trials. The mean error rate for target-present trials was 10.3%. Error rates were analysed with the same planned comparisons reported in Table 3 for the search latencies. Based on these results, we can highlight the following observations. (1) The comparison 8a1h–8h1a was significant, indicating a larger number of errors for a happy target face in an angry crowd than the opposite, $z = 2.84$, $p = .025$. (2) The comparison 8a1n–8n1a was significant, indicating a larger number of errors for a neutral target face in an angry crowd than the opposite, $z = 3.87$, $p = .001$. (3) None of the other planned comparisons were statistically significant.

In a separate analysis, we computed the error rates as a function of target valence (happy $= 12.1\%$, neutral $= 12.2\%$, angry $= 6.9\%$). Multiple comparisons with Tukey contrasts indicated that (1) the error rate was larger for happy than for angry target faces, $z = 3.03$, $p = .0069$, (2) the error rate was

---

[7] The p-values for the t-test are based on the Satterthwaite approximation for the degrees of freedom.

larger for neutral than for angry target faces, $z = 2.92$, $p = .010$, and (3) the error rates did not differ across happy and neutral target faces, $z = -0.12$. *Target-absent trials.*   The mean error rate for target-absent trials was 2.4%. The error rate was higher for angry crowds (3.2%) than for happy crowds (0.7%), $z = 2.70$, $p = .017$. The error rate was higher for neutral crowds (3.4%) than for happy crowds, $z = 2.53$, $p = .028$. There was no statistically significant difference between the error rates of angry and neutral crowds, $z = 0.05$.

Discussion

The results of Experiment 2 indicate that (1) an angry (dynamic) face is more efficiently detected in a crowd of happy faces than the opposite, (2) search times are longer in an all-angry crowd than in an all-happy crowd, and (3) angry faces are more efficiently detected in neutral crowds than are neutral faces in angry crowds. Whereas all these three results are consistent with the ASE, the third result is inconsistent with the "enhanced dwell-time" hypothesis (i.e., the delayed disengagement from threatening faces).

   The results reported in Figure 7 show that angry faces in neutral crowds are detected more efficiently than are happy faces in neutral crowds. When considering the emotional crowds, moreover, a facilitation for singleton targets is found for the angry expression relative to the neutral expression 8h1a–8h1n, but not for the happy expression relative to the neutral expression 8a1h–8a1n. These results suggest that dynamic angry faces are more efficiently detected relative to dynamic happy or neutral faces.

EXPERIMENTS 3A AND 3B

Images of faces that have an open mouth relative to close-mouthed faces yield a simple visual feature (visible teeth) that can drive efficient search because it easily differentiates the target from the distractors (Becker, Anderson, et al., 2011; Horstmann et al., 2012; Purcell et al., 1996). Rather than trying to eliminate any spurious difference in bottom-up saliency between happy and angry faces, in Experiments 3A and 3B we replicated the design of Experiments 1A and 1B, but we purposely used open-mouthed happy faces and closed-mouthed angry faces, so as to favour the HSE. In fact, it is easier to detect a more salient target among less salient distractors than the opposite (Wolfe, 2001). In these conditions, if the HSE is found, we cannot conclude whether it is due to the stimulus-driven factors or to the expressive properties of the faces. On the other hand, if we observe an ASE, we should conclude that it is due to the expressive content of the faces, because the bottom-up saliency predicts the opposite result.

Method

Participants.   A total of 26 undergraduate students from Florence University participated in the experiment, 16 women and 10 men, with a mean age of 28 years ($SD = 4.2$). Twelve undergraduates participated in Experiment 3A and 14 undergraduates participated in Experiment 3B. All participants were naïve to the purpose of the study. All participants had normal or corrected-to-normal vision. Participation was voluntary. None of them had participated in the previous experiments.

Apparatus.   The apparatus was the same as in the previous experiments.

Stimuli.   We used the same nine face identities selected from the Radboud database as in Experiments 1A and 1B. Differently than in the previous experiments, however, the pictures of the happy faces displayed an enlarged open mouth with visible teeth (see Figure 1c).
Amount of movement.   The nine face identities produced similar amounts of motion in the dynamic condition when displaying happy or angry expressions, $t(15.34) = 0.25$, $p = .81$ (Welch correction).
Bottom-up visual saliency.   To evaluate the bottom-up saliency, we generated 1600 displays comprising one target face embedded in a crowd of seven face distractors. As in Experiments 1A and 1B, these displays were processed with the SaliencyToolbox 2.2 (Walther & Koch, 2006). Differently than in Experiments 1A and 1B, the happy target faces were 11% more salient than the angry target faces, $t(1591) = 13.34$, $p = .001$.

Design and procedure.   The design and procedure of Experiment 3A were the same as in Experiments 1A. The design and procedure of Experiment 3B were the same as in Experiments 1B.

Results—Experiment 3A

Response latencies.   Mean detection latency on "correct detection" trials as a function of Crowd, stimulus, and target presence is shown in Figure 8. On average, participants responded 619 ms faster in target-present than in target-absent trials, $\chi_1^2 = 23.7$, $p = .001$.
Target-present trials.   The Stimulus × Crowd interaction was statistically significant, $\chi_1^2 = 7.22$, $p = .007$. Under dynamic conditions, average RTs on "correct detection" trials were 268 ms faster for angry-face targets among happy-face distractors than the opposite (happy-face targets among angry-face distractors), $\chi_1^2 = 4.50$, $p = .034$. Under static conditions, the search for an angry face within a crowd of happy faces did not produce RTs
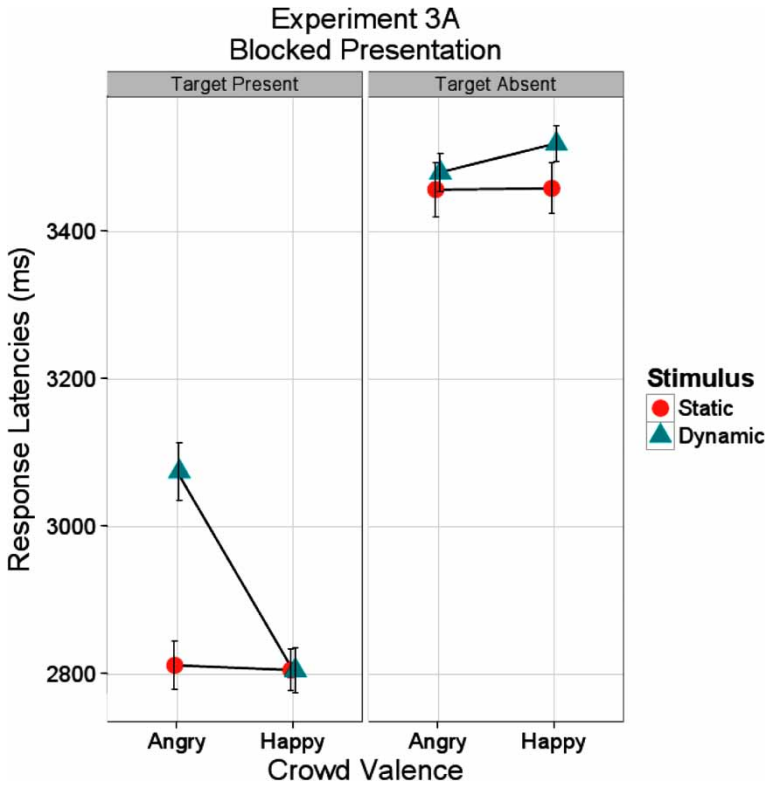
Figure 8.   Experiment 3A. Response latencies as a function of target (present, absent), crowd valence (angry, happy), and stimulus (circle: static; triangle: dynamic). Error bars represent ±one standard error. To view this figure in colour, please see the online issue of the Journal.

significantly different from those of the search for a happy face within a crowd of angry faces, $\chi_1^2 = 0.03$, $p = .86$.

Target-absent trials.   There were no significant effects for crowd, stimulus, or their interaction, $\chi_3^2 = 1.54$, $p = .67$.

Error analysis.   Mean percentage errors are presented in Table 5 as a function of target (present, absent), stimulus (static, dynamic), and crowd valence (happy, angry).

Target-present trials.   The mean error rate for target-present trials was 3.6%. There were no significant effects for crowd, stimulus, or their interaction, $\chi_7^2 = 0.26$, $p = .97$.

Target-absent trials.   The mean error rate for target-absent trials was less than 1%.

TABLE 5
Mean percentage errors for target-present and target-absent trials in the visual search task of Experiment 3A as a function of stimulus (static, dynamic) and crowd valence (angry, happy)

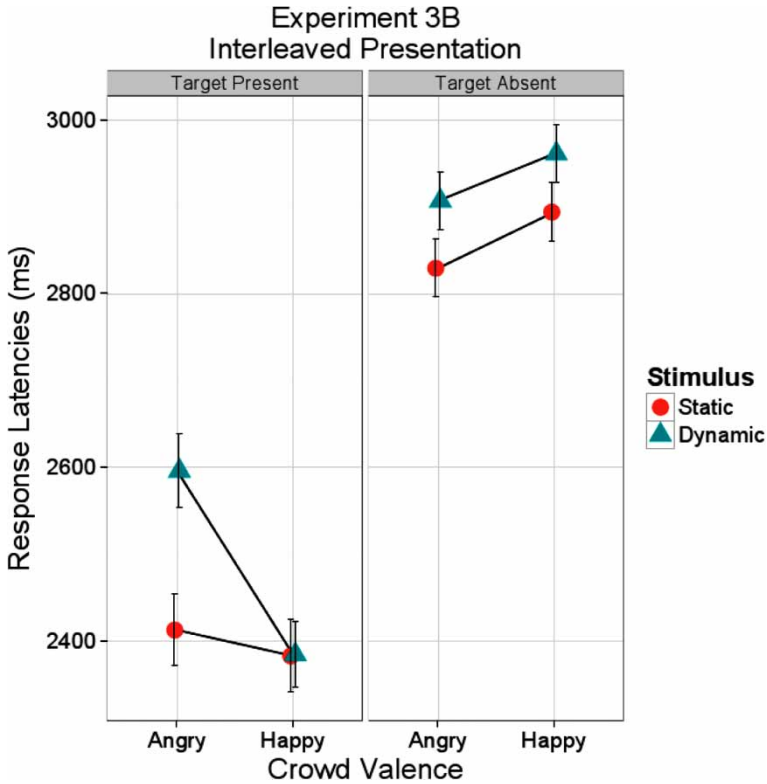| Target	Stimulus	Crowd valence	% errors
Present	Static	Angry	3.8
		Happy	3.3
	Dynamic	Angry	4.1
		Happy	3.4
Absent	Static	Angry	0.0
		Happy	0.4
	Dynamic	Angry	0.4
		Happy	0.0

Figure 9. Experiment 3B. Response latencies as a function of target (present, absent), crowd valence (angry, happy), and stimulus (circle: static; triangle: dynamic). Error bars represent ±one standard error. To view this figure in colour, please see the online issue of the Journal.

TABLE 6

Mean percentage errors for target-present and target-absent trials in the visual search task of Experiment 3B as a function of stimulus (static, dynamic) and crowd valence (angry, happy)

Target	Stimulus	Crowd valence	% errors
Present	Static	Angry	10.9
		Happy	11.8
	Dynamic	Angry	12.4
		Happy	11.1
Absent	Static	Angry	5.6
		Happy	0.5
	Dynamic	Angry	6.4
		Happy	0.9

Results—Experiment 3B

Mean detection latency on correct trials as a function of crowd valence (happy, angry), target (present, absent), and stimulus (static, dynamic) is shown in Figure 9. On average, participants responded 454 ms faster in target-present than in target-absent trials, $\chi^2_1 = 27.6$, $p = .001$.

Target-present trials. The Stimulus × Crowd interaction was statistically significant, $\chi^2_1 = 5.25$, $p = .022$. Under dynamic conditions, average RTs on "correct detection" trials were 197 ms faster for angry-face targets among happy-face distractors than the opposite (happy-face targets among angry-face distractors), $\chi^2_1 = 7.72$, $p = .0055$. Under static conditions, the search for an angry face within a crowd of happy faces did not produce RTs significantly different from those of the search for a happy face within a crowd of angry faces, $\chi^2_1 = 0.17$, $p = .68$.

Target-absent trials. There were no significant effects for crowd, stimulus, or their interaction, $\chi^2_3 = 3.95$, $p = .27$.

Error analysis. Mean percentage errors are presented in Table 6 as a function of target (present, absent), stimulus (static, dynamic), and crowd valence (happy, angry).

Target-present trials. The mean error rate for target-present trials was 11.6%. There were no significant effects for crowd, stimulus, or their interaction, $\chi^2_3 = 1.05$, $p = .79$.

Target-absent trials. The mean error rate for target-absent trials was 3.3%. The error rate for the angry crowds was 5.2% larger than for happy crowds, $\chi^2_1 = 11.0$, $p = .001$. No other effects or interactions were statistically significant.

Comparison of Experiments 3A and 3B

Response latencies.   The mean response latencies on "correct detection" trials were longer in Experiment 3A (blocked presentation) than in Experiment 3B (interleaved presentation), both on target-present trials (2879 ms vs. 2460 ms), $\chi^2_1 = 4.19$, $p = .041$, and target-absent trials (3494 ms vs. 2926 ms), $\chi^2_1 = 9.9$, $p = .0017$.[8]

Error analysis.   Consistently with the pattern of the RTs, the error rates were smaller in Experiment 3A than in Experiment 3B, both on target-present trials (3.6% vs. 11.3%), $\chi^2_1 = 4.99$, $p = .026$, and target-absent trials (0.1% vs. 3.1%), $p = .0022$, $\chi^2_1 = 9.38$

Discussion

Open-mouthed smiling faces provide a high-contrast, low-level feature that could drive efficient search. In fact, an analysis of our stimuli based on the Itti-Koch model (Itti & Koch, 2000, 2001; Koch & Ullman, 1985) confirmed that happy faces were more salient than angry faces. If search latencies mainly reflect bottom-up saliency processes, then we should find a search asymmetry favouring happy faces. Instead, under dynamic conditions (but not under static conditions), we found a search asymmetry favouring angry faces, in both Experiments 3A (blocked presentation) and 3B (interleaved presentation).

In the target-absent trials, the mean search latencies were not modulated neither by the affective valence of the distractors nor by the type of presentation (static, dynamic). These results weaken the hypothesis that the ASE found under dynamic conditions might due to a "delayed disengagement" from threatening faces.

As a final caveat, we must add that, even though the target-absent decisions for angry and happy crowds were no different, it is still possible that participants process happy faces more efficiently than angry faces, and that allows the dynamic anger to be better detected. There might be a number of alternative explanations for why the target-absent decisions are equivalent in the present experiments. One possibility is that the happy crowds are efficiently searched, but given the costs of missing anger, participants are being more cautious about saying there is no angry target. The accuracy results are consistent with this possibility.

---

[8] As for Experiments 1A and 1B, this result can be attributed to the different stimulus exposure times in the two experiments.

GENERAL DISCUSSION

Visual search asymmetries in tasks involving emotional faces may be attributed to two factors: (1) bottom-up factors that make some faces in a crowd to be more visually distinctive that others, without reflecting the recognition of an emotion (e.g., Calvo & Nummenmaa, 2008), and (2) a preferential top-down orienting of attention, either towards danger-related stimuli (e.g., Pourtois, Grandjean, Sander, & Vuilleumier, 2004) or towards happy faces (e.g., Brosch, Sander, Pourtois, & Scherer, 2008). To our knowledge, this is the first study in which the face-in-the-crowd effect is tested by using dynamic images of realistic human faces, with validated emotional expressions having similar intensities, after controlling the bottom-up visual saliency and the amount of image motion. Our purpose was to determine whether, under static and/or dynamic conditions, such stimuli induce a visual search asymmetry that cannot be explained by bottom-up factors.

In Experiments 1A and 1B, we used closed mouth happy faces and closed mouth angry faces. In Experiment 2 we included neutral faces as a baseline condition. The stimuli were selected so that the bottom-up factors did not favour either the ASE or the HSE. In Experiments 3A and 3B, we used open mouth happy faces and closed mouth angry faces. In these conditions, the bottom-up saliency, as determined by the computational model of attention of Itti and Koch (2000, 2001), favoured the HSE. Regardless of the low-level or stimulus-driven factors, the results of all the present experiments provide evidence for an ASE when the emotions unfold in a dynamic fashion, but no evidence for either the ASE or the HSE when the facial emotional expressions are presented as static images.

We propose to explain the different results obtained under static and dynamic conditions in the following manner. In Experiments 1A, 1B, and 2, the happy and angry faces were matched for bottom-up visual saliency. We speculate that, under static conditions, the top-down orienting bias towards threatening faces was too weak to generate an ASE. Instead, under dynamic conditions the top-down bias was enhanced and an ASE was observed.

In Experiments 3A and 3B, the bottom-up factors favoured the HSE, whereas the top-down factors favoured the ASE (if they correspond to an orienting bias towards threatening faces). We speculate that, under static conditions, these two factors balanced each other and no search asymmetry emerged. However, the top-down factors might be enhanced in response to dynamic facial expressions of emotion, and this might have shifted the balance towards the ASE. In fact, evidence from both behavioural (Sato & Yoshikawa, 2007a, 2007b; Yoshikawa & Sato, 2008) and neuroimaging studies (Kilts, Egan, Gideon, Ely, & Hoffman, 2003; LaBar et al., 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Schultz & Pilz, 2009;

Trautmann, Fehr, & Herrmann, 2009) indicates that dynamic facial expressions induce stronger responses than static expressions.

It has been suggested that angry faces might be more effective in engaging visual attention than other valenced faces, and this might explain the ASE (Fox et al., 2002). However, our results are inconsistent with the idea that the ASE depends on a delayed disengagement from threatening faces. In fact, in Experiment 2 we found that a dynamic angry singleton face is more rapidly detected in a crowd of neutral faces than the opposite (i.e., a neutral singleton face among angry faces).

The use of a neutral baseline in Experiment 2 also suggests that dynamic angry faces might be detected more efficiently than dynamic friendly faces. In fact, participants were faster to find an angry face among neutral faces than to find a happy face among neutral faces (but see the cautionary note in the Discussion section of Experiments 3A and 3B).

The results obtained in the present experiments with static faces replicate previous findings showing that the ASE does not emerge when using photographs of realistic faces with different identities, which reproduce the natural variability of emotional expressions (Becker, Anderson et al., 2011; Becker et al., 2012; Calvo & Nummenmaa, 2008; Juth et al., 2005; Lipp, Price, & Tellegen, 2009a, 2009b; Purcell et al., 1996). The novel result of the present study is that the same face identities can generate an ASE, if their facial expressions are displayed dynamically.

It is important to point out that our dynamic stimuli do not replicate the temporal dynamics of natural facial behaviour. The image sequences used in the dynamic condition displayed morph sequences which changed the face in a linear fashion over time from neutral to a peak expression. Moreover, the photographs used to generate the morph sequences were selected so as to produce similar amounts of image deformation across facial expressions (see Appendix A). However, even under such gross approximations, the dynamic displays provided a greater degree of realism than the static displays, with strong behavioural consequences (see Figures 4, 5, 7, 8, and 9). In fact, this method for generating dynamic emotional expressions has been successfully used in many previous behavioural studies (e.g., Becker et al., 2012; Joormann & Gotlib, 2006; Kamachi et al., 2001; Lynch et al., 2006; Marinetti, Mesquita, Yik, Cragwall, & Gallagher, 2012; Oosterhof & Todorov, 2009; Sato & Yoshikawa, 2004; Yoshikawa & Sato, 2008) and it has been shown to elicit an enhanced neural response with respect to static emotional facial expressions (Sato, Fujimura, & Suzuki, 2008; Sato et al., 2004; see also LaBar et al., 2003). Interestingly, Ambadar, Schooler, and Cohn (2005) measured the accuracy of emotion recognition from video-ecordings that preserved the dynamic characteristics of natural facial expressions of basic emotions ("dynamic" condition) and from frame sequences in which only the first and last frame of the original video

sequences were presented ("first-last" condition). In two control conditions, a 200-ms mask was inserted between successive images of the original video sequence ("multistatic" condition), or only the last frame of the video sequence was presented ("single-static" condition). Ambadar et al. found that the dynamic and first-last condition resulted in improved recognition accuracy over the single-static and multistatic conditions. Importantly, no differences were found in accuracy of recognition between the dynamic and the first-last conditions. Therefore, Ambadar et al. concluded that the temporal dynamics of natural facial behaviour provide no further processing advantages beyond those elicited by motion information generated in the first-last condition.

The use of realistic images portraying different face identities is important for ecological considerations. Nevertheless, Horstmann et al. (2012) pointed out that, often, different results are obtained with images selected from different databases. Stimuli selected from the Pictures of Facial Affect (Ekman & Friesen, 1975) are more likely to elicit an ASE (e.g., Fox & Damjanovic, 2006) than the HSE (but see Becker, Anderson et al., 2011); conversely, the HSE has been observed more often for stimuli selected from the Karolinska Directed Emotional Faces (KDEF; Calvo & Nummenmaa, 2008; Juth et al., 2005; Lundqvist, Flykt, & Ohman, 1998) and the MacBrain stimulus set (or "NimStim"; Becker, Anderson et al., 2011; Horstmann et al., 2012; Tottenham et al., 2009; Williams, Moss, Bradshaw, & Mattingley, 2005). These differences may depend on a number of factors. In a visual search paradigm, attention is guided by both bottom-up and top-down factors (see Becker, 2010). (1) Stimulus-driven factors depend on the similarity between target and distractors, and on the dissimilarity among the distractors (e.g., Guest & Lamberts, 2011). In this respect, it is reasonable to assume that the relative bottom-up saliency of face images with different expressions and identities varies across multiple databases. To our knowledge, the HSE has never been found by using face images selected from the Radboud Face Database. (2) Top-down control, which tunes attention towards specific features and favours the selection of those items that match the feature values of the search target, also affects the ASE. Top-down control might have stronger or weaker effects depending on the experimental design: for example, when participants are instructed to search for a specific target among known distractors (e.g., Horstmann et al., 2012) or when, in each trial, participants do not know beforehand which face expression will be used as target and which face expression will be used for the distractors (as in the present case).

It is interesting to compare the present results with those of Pinkham et al. (2010). Pinkham et al. used a visual search paradigm in which the targets were photographs of faces with validated facial expressions within a

crowd of nonrepeating facial identities. Therefore, their stimuli are similar to those presently used in the static condition.

However, there are also differences between our stimuli and theirs, specifically, in the database from which the images were selected, in the rendering of the faces (grayscale vs. colour), in the control of the emotional intensity, and in the control of the bottom-up saliency. Perhaps because of these differences, an ASE was found with the stimuli of Pinkham et al. (2010) but not with our static stimuli.

The discrepant results from these two studies, as well as other inconsistencies in the literature, indicate that, with static facial displays of emotion, the presence/absence of an ASE depends on the specific features of the stimulus set and on the experimental design. It remains an interesting question for future research to determine whether *dynamic* facial displays of emotion, under different experimental designs and with stimulus images selected from multiple databases, might provide a stronger support to the hypothesis that a threat detection mechanism is triggered when searching for an angry face in a crowd.

The present investigation is not without limitations. First, we only used one set size. This would be a serious limitation, had the goal been to demonstrate that the detection of an angry face in a crowd of face distractors occurs through parallel, rather than serial, searching. However, the literature overwhelmingly indicates that, when searching for a face in a crowd, the search that is performed is a serial one (Becker, Anderson et al., 2011). Therefore, our goal was not to distinguish between serial and parallel processing, but rather to investigate search asymmetries among affective faces. Second, we did not test a face-inversion condition. Although face inversion provides a control condition in many circumstances, no clearcut predictions can be made for emotional faces. It is well established that face inversion reduces holistic face-processing (Caudek & Lorenzino, 2012; Rossion, 2009; Rossion & Gauthier, 2002; Yin, 1969), but there is also evidence that the emotional expression of faces can still be recognized well above chance level when faces are presented upside down (e.g., McKelvie, 1995). So, it is not entirely obvious that a search asymmetry favouring angry faces should disappear under conditions of face inversion (Becker, Anderson, et al., 2011). Third, our stimuli always displayed whole faces. Therefore, we cannot exclude that specific facial features may have contributed to the observed results (Pinkham et al., 2010; Schubö et al., 2006). In the case of static photorealistic emotional faces, Horstmann et al. (2012) showed that the ASE is in a large measure due to the presence of visible teeth. Fox and Damjanovic (2006) showed that the eyes, even when presented in isolation, can produce the ASE. The current experiments address the saliency of whole faces, but not the salience of the individual parts of faces. Until that is done, we do not know if the ASE is due to the processing of the whole face, or to

the tendency to pay more attention to some components of the face. Fourth, in Experiments 1A, 1B, 3A, and 3B, a search-asymmetry design was used, even if it has been argued that the constant distractor paradigm is better suited to study the differential guidance of focal attention (Becker, Anderson, et al., 2011; Eastwood et al., 2001; Frischen et al., 2008). We chose a search-asymmetry design because there is not a unique way to create nonexpressive (distractor) dynamic faces. In Experiment 2, we dealt with this problem by displaying the movements of the mouth and face when pronouncing a phoneme, but other alternatives are possible (e.g., chewing, yawning, and so on). However, whatever choice is made it does not guarantee that neutral and happy faces, on the one side, and neutral and angry faces, on the other, are separated by the same distance within the psychological similarity space (Horstmann, Scharlau, & Ansorge, 2006; Krysko & Rutherford, 2009). Fifth, there are other negative emotions, besides anger (e.g., fear, sadness, disgust, and so on). Williams et al. (2005) proposed that angry and fearful faces might convey two different kinds of threat: An angry face signals that a particular individual is the locus of the threat, whereas a fearful face signals that the threat is located elsewhere in the environment. It would be interesting to compare search efficiency for faces displaying different kinds of negative emotions, under static and dynamic conditions.

Our results are consistent with the hypothesis that angry faces can capture attention in an involuntary, bottom-up manner. It has been shown, however, that these involuntary shifts in attention may be modulated, and also overridden, by voluntary recruitment of attention (Mohanty & Sussman, 2013). For example, Hahn and Gronlund (2007) found that threatening facial expressions are searched more efficiently than happy facial expressions when they are the target of the search, but not when they are in opposition to the task goal. In one of their experiments, participants were instructed to search for a specific facial expression (i.e., either an angry face or a happy face in a neutral crowd). In some trials, a nontarget singleton was shown (e.g., a happy face in a neutral crowd, when the task was to search for an angry face). For these nontarget singletons, which were in opposition to the task goal, Hahn and Gronlund found equivalent search slopes for angry and happy faces. Therefore, their results suggest that, in addition to the involuntary capture of attention by emotional stimuli, visual search is also affected by top-down (emotional/motivational) factors (see also Horstmann & Becker, 2008).

Notwithstanding the limitations of this study, our results yield evidence of a search asymmetry favouring threatening over friendly faces, also with ecologically valid stimuli. Our results indicate that (1) emotional expressions can overcome the effects of perceptual factors and can modulate the face-in-the-crowd effect in the direction of the ASE, and (2) dynamic facial information can reveal search asymmetries that may go unnoticed when using static faces.

REFERENCES

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16*, 403–410. doi:10.1111/j.0956-7976.2005.01548.x

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain*, *24*, 149–163.

Baayen, R. H. (2008). *Practical data analysis for the language sciences with R*. Cambridge, MA: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*, 12–28.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., & Dai, B. (2008). *lme4: Linear mixed-effects models using S4 classes* (R package version 0.999375-28) (Computer software manual). Retrieved from http://lme4.r-forge.r-project.org

Batty, M. J., Cave, K. R., & Pauli, P. (2005). Abstract stimuli associated with threat through conditioning cannot be detected preattentively. *Emotion*, *5*, 418–430. doi:10.1037/1528-3542.5.4.418

Becker, S. I. (2010). The role of target-distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology: General*, *139*, 247–265.. doi:10.1037/a0018808

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S., & Neel, R. (2011). The face in the crowd effect unconfounded: Happy faces, not angry faces, are more efficiently detected in the visual search task. *Journal of Experimental Psychology: General*, *140*, 637–659. doi:10.1037/a0024060

Becker, D. V., Neel, R., Srinivasan, N., Neufeld, S., & Kumar, D. (2012). The vividness of happiness in dynamic facial displays of emotion. *PLoS ONE*, *7*(1), e26551. doi:10.1371/journal.pone.0026551

Becker, S. I., Horstmann, G., & Remington, R. W. (2011). Perceptual grouping, not emotion, accounts for search asymmetries with schematic faces. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1739–1757. doi:10.1037/a0024665

Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: Rapid Spatial orienting toward positive emotional stimuli. *Psychological Science*, *19*, 362–370. doi:10.1111/j.1467-9280.2008.02094.x

Calvo, M. G., Avero, P., & Lundqvist, D. (2006). Facilitated detection of angry faces: Initial orienting and processing efficiency. *Cognition and Emotion*, *20*, 785–811. doi:10.1080/02699930500465224

Calvo, M. G., & Esteves, F. (2005). Detection of emotional faces: Low perceptual threshold and wide attentional span. *Visual Cognition*, *12*, 13–27. doi:10.1080/13506280444000094

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: Salient physical features guide effective visual search. *Journal of Experimental Psychology: General*, *137*, 471–494. doi:10.1037/a0012771

Caudek, C., & Domini, F. (2013). Priming effects under correct change detection and change blindness. *Consciousness and Cognition*, *22*, 290–305. doi:10.1016/j.concog.2012.08.003

Caudek, C., & Lorenzino, M. (2012, May). *Recognition memory is more accurate when faces are inverted than when they are upright*. Poster presented at the annual meeting of the Vision Sciences Society, Naples, FL.

Caudek, C., & Monni, A. (2013). Do you remember your sad face? The roles of negative cognitive style and sad mood. *Memory*, 1–13. doi:10.1080/09658211.2013.765893

Cerf, M., Harel, J., Einhäuser, W., & Koch, C. (2008). Predicting human gaze using low-level saliency combined with face detection. *Methods*, *20*, 241–248.

Chentsova-Dutton, Y. E., & Tsai, J. L. (2010). Self-focused attention and emotional reactivity: The role of culture. *Journal of Personality and Social Psychology*, *98*, 507–519. doi:10.1037/a0018534

Coelho, C. M., Cloete, S., & Wallis, G. (2010). The face-in-the-crowd effect: When angry faces are just cross(es). *Journal of Vision*, *10*, 1–14.

Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*, *20*, 125–151. doi:10.1016/0278-2626(92)90065-T

den Uyl, M. J. D., & van Kuilenburg, H. (2005). The FaceReader: Online facial expression recognition. In *Proceedings of measuring behavior 2005* (pp. 589–590). Wageningen, The Netherlands.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception and Psychophysics*, *63*, 1004–1013. doi:10.3758/BF03194519

Ekman, P., & Friesen, W. V. (1975). Unmasking the face: A guide to recognizing emotions from facial clues. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system: Manual*. Salt Lake City, UT: Research Nexus.

Foley, E., Rippon, G., Thai, N. J., Longe, O., & Senior, C. (2012). Dynamic facial expressions evoke distinct activation in the face perception network: A connectivity analysis study. *Journal of Cognitive Neuroscience*, *24*, 507–520. doi:10.1162/jocn_a_00120

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*, 6–16. doi:10.1167/8.2.6

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion*, *6*, 534–539. doi:10.1037/1528-3542.6.3.534

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion*, *14*, 61–92. doi:10.1080/026999300378996

Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition and Emotion*, *16*, 355–379. doi:10.1080/02699930143000527

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological Bulletin*, *134*, 662–676. doi:10.1037/0033-2909.134.5.662

Fujimura, T., & Suzuki, N. (2010). Effects of dynamic information in recognising facial expressions on dimensional and categorical judgements. *Perception*, *39*, 543–552. doi:10.1068/p6257

Fusar-Poli, P., Placentino, A., Carletti, F., Allen, P., Landi, P., Abbamonte, M., . . . Politi, P. L. (2009). Laterality effect on emotional faces processing: ALE meta-analysis of evidence. *Neuroscience Letters*, *452*, 262–267. doi:10.1016/j.neulet.2009.01.065

Guest, D., & Lamberts, K. (2011). The time course of similarity effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1667–1688. doi:10.1037/a0025640

Hahn, S., & Gronlund, S. (2007). Top-down guidance in visual search. *Psychonomic Bulletin and Review*, *14*, 159–165. doi:10.3758/BF03194044

Hansen, C., & Hansen, R. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology*, *54*, 917–924. doi:10.1037/0022-3514.54.6.917

Horstmann, G. (2007). Preattentive face processing: What do visual search experiments with schematic faces tell us? *Visual Cognition*, *15*, 799–833. doi:10.1080/13506280600892798

Horstmann, G. (2009). Visual search for schematic affective faces: Stability and variability of search slopes with different instances. *Cognition and Emotion*, *23*, 355–379. doi:10.1080/02699930801976523

Horstmann, G., & Ansorge, U. (2009). Visual search for facial expressions of emotions: A comparison of dynamic and static faces. *Emotion*, *9*, 29–38. doi:10.1037/a0014147

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion*, *6*, 193–207. doi:10.1037/1528-3542.6.2.193

Horstmann, G., & Becker, S. I. (2008). Attentional effects of negative faces: Top-down contingent or involuntary? *Perception and Psychophysics*, *70*, 1416–1434. doi:10.3758/PP.70.8.1416

Horstmann, G., Lipp, O. V., & Becker, S. I. (2012). Of toothy grins and angry snarls: Open mouth displays contribute to efficiency gains in search for emotional faces. *Journal of Vision*, *12*, 1–15. doi:10.1167/12.5.7

Horstmann, G., Scharlau, I., & Ansorge, U. (2006). More efficient rejection of happy than angry face distractors in visual search. *Psychonomic Bulletin and Review*, *13*, 1067–1073. doi:10.3758/BF03213927

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*, 346–363. doi:10.1002/bimj.200810425

Humphrey, K., Underwood, G., & Lambert, T. (2012). Salience of lambs: A test of the saliency map hypothesis with pictures of emotive objects. *Journal of Vision*, *12*, 1–15. doi:10.1167/12.1.22

Hunt, A. R., Cooper, R. M., Hungr, C., & Kingstone, A. (2007). The effect of emotional faces on eye movements and attention. *Visual Cognition*, *15*, 513–531. doi:10.1080/13506280600843346

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506. doi:10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–203. doi:10.1038/35058500

Joormann, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology*, *115*, 705–714. doi:10.1037/0021-843X.115.4.705

Judd, C., Westfall, J., & Kenny, D. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54–69. doi:10.1037/a0028347

Juth, P., Lundqvist, D., Karlsson, A., & Öhman, A. (2005). Looking for foes and friends: Perceptual and emotional factors when finding a face in the crowd. *Emotion*, *5*, 379–395. doi:10.1037/1528-3542.5.4.379

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, *30*, 875–887. doi:10.1068/p3131

Kaulard, K., Cunningham, D. W., Bülthoff, H. H., & Wallraven, C. (2012). The MPI Facial Expression Database: A validated database of emotional and conversational facial expressions. *PLoS ONE*, *7*, e32321. doi:10.1371/journal.pone.0032321

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage*, *18*, 156–168. doi:10.1006/nimg.2002.1323

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Psychology*, *1*(January), 12.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.

Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research*, *51*, 1457–1483. doi:10.1016/j.visres.2010.12.014

Krysko, K. M., & Rutherford, M. D. (2009). The face in the crowd effect: Threat-detection advantage with perceptually intermediate distractors. *Visual Cognition*, *17*, 1205–1217. doi:10.1080/13506280902767789

LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, *13*, 1023–1033. doi:10.1093/cercor/13.10.1023

Lander, K., & Chuang, L. (2005). Why are moving faces easier to recognize? *Visual Cognition*, *12*, 429–442. doi:10.1080/13506280444000382

Lander, K., Chuang, L., & Wickham, L. (2006). Recognizing face identity from natural and morphed smiles. *Quarterly Journal of Experimental Psychology*, *59*, 801–808. doi:10.1080/17470210600576136

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*, 1377–1388. doi:10.1080/02699930903485076

Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009a). Emotional faces in neutral crowds: Detecting displays of anger, happiness, and sadness on schematic and photographic images of faces. *Emotion*, *33*, 240–260.

Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009b). No effect of inversion on attentional and affective processing of facial expressions. *Emotion*, *33*, 248–259. doi:10.1037/a0014715

Lynch, T. R., Rosenthal, M. Z., Kosson, D. S., Cheavens, J. S., Lejuez, C. W., & Blair, R. J. R. (2006). Heightened sensitivity to facial expressions of emotion in borderline personality disorder. *Emotion*, *6*, 647–655. doi:10.1037/1528-3542.6.4.647

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces*. Stockholm: Karolinska Institute, Psychology Section, Department of Clinical Neuroscience.

Lundqvist, D., & Öhman, A. (2005). Emotion regulates attention: The relation between facial configurations, facial emotion and visual attention. *Visual Cognition*, *12*, 51–84. doi:10.1080/13506280444000085

Mak-Fan, K. M., Thompson, W. F., & Green, R. E. A. (2011). Visual search for schematic emotional faces risks perceptual confound. *Cognition and Emotion*, *4*, 573–584.

Mannan, S. K., Kennard, C., & Husain, M. (2009). The role of visual salience in directing eye movements in visual object agnosia. *Current Biology*, *19*, R247–R248. doi:10.1016/j.cub.2009.02.020

Marinetti, C., Mesquita, B., Yik, M., Cragwall, C., & Gallagher, A. H. (2012). Threat advantage: Perception of angry and happy dynamic faces across cultures. *Cognition and Emotion*, *26*, 1326–1334. doi:10.1080/02699931.2011.644976

Mather, M., & Knight, M. R. (2006). Angry faces get noticed quickly: Threat detection is not impaired among older adults. *Journals of Gerontology: Psychological Sciences and Social Sciences*, *61B*, P54–P57. doi:10.1093/geronb/61.1.P54

McKelvie, S. I. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology*, *34*, 325–334. doi:10.1111/j.2044-8309.1995.tb01067.x

Mermillod, M., Vermeulen, N., Lundqvist, D., & Niedenthal, P. M. (2009). Neural computation as a tool to differentiate perceptual from emotional processes: The case of anger superiority effect. *Cognition*, *110*, 346–357. doi:10.1016/j.cognition.2008.11.009

Mohanty, A., & Sussman, T. J. (2013). Top-down modulation of attention by emotion. *Frontiers in Human Neuroscience*, *7*, 1–5. doi:10.3389/fnhum.2013.00102

Musterle, W., & Rossler, O. E. (1986). Computer faces: The human Lorenz matrix. *Bio Systems*, *19*, 61–80. doi:10.1016/0303-2647(86)90035-3

Noldus Information Technologies. (2012). *FaceReader 4.0: Technical specifications*. Leesburg, VA: Author.

Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision*, *8*, 1–23. doi:10.1167/8.8.1

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. doi:10.1037/0022-3514.80.3.381

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*, 128–133. doi:10.1037/a0014520

O'Toole, A., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., & Dunlop, J. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*, 74–83. doi:10.1016/j.visres.2010.09.035

O'Toole, A., Roark, D., & Abdi, H. (2002). Recognition of moving faces: A psychological and neural perspective. *Trends in Cognitive Science*, *6*, 261–266. doi:10.1016/S1364-6613(02)01908-3

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123. doi:10.1016/S0042-6989(01)00250-4

Pelphrey, K. A., Morris, J. P., McCarthy, G., & LaBar, K. S. (2007). Perception of dynamic changes in facial affect and identity in autism. *Social Cognitive and Affective Neuroscience*, *2*, 140–149. doi:10.1093/scan/nsm010

Pilz, K. S., Bülthoff, H. H., & Vuong, Q. C. (2009). Learning influences the encoding of static and dynamic faces and their recognition across different spatial frequencies. *Visual Cognition*, *17*, 716–735. doi:10.1080/13506280802340588

Pilz, K. S., Thornton, I. M., & Bülthoff, H. H. (2006). A search advantage for faces learned in motion. *Experimental Brain Research*, *171*, 436–447. doi:10.1007/s00221-005-0283-8

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-PLUS. In J. Chambers, W. Eddy, W. Härdle, S. Sheather, & L. Tierney (Eds.), *Statistics and computing* (pp. 57–96). New York, NY: Springer-Verlag.

Pinkham, A. E., Griffin, M., Baron, R., Gur, R. C., & Sasson, N. J. (2010). The face in the crowd effect: Anger superiority when using real faces and multiple identities. *Emotion*, *10*, 141–146. doi:10.1037/a0017387

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex*, *14*, 619–633. doi:10.1093/cercor/bhh023

Purcell, D. G., & Stewart, A. L. (2010). Still another confounded face in the crowd. *Attention, Perception, and Psychophysics*, *72*, 2115–2127.

Purcell, D. G., Stewart, A. L., & Skov, R. B. (1996). It takes a confounded face to pop out of a crowd. *Perception*, *25*, 1091–1108. doi:10.1068/p251091

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica*, *132*, 300–312. doi:10.1016/j.actpsy.2009.08.002

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*, 63–75. doi:10.1177/1534582302001001004

Sato, W., Fujimura, T., & Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. *International Journal of Psychophysiology*, *70*, 70–74. doi:10.1016/j.ijpsycho.2008.06.001

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Cognitive Brain Research*, *20*, 81–91. doi:10.1016/j.cogbrainres.2004.01.008

Sato, W., & Yoshikawa, S. (2004). The dynamic aspects of emotional facial expressions. *Cognition and Emotion*, *18*, 701–710. doi:10.1080/02699930341000176

Sato, W., & Yoshikawa, S. (2007a). Enhanced experience of emotional arousal in response to dynamic facial expressions. *Journal of Nonverbal Behavior*, *31*, 119–135. doi:10.1007/s10919-007-0025-7

Sato, W., & Yoshikawa, S. (2007b). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, *104*, 1–18. doi:10.1016/j.cognition.2006.05.001

Schubö, A., Gendolla, G. H. E., Meinecke, C., & Abele, A. E. (2006). Detecting emotional faces and features in a visual search paradigm: Are faces special? *Emotion*, *6*, 246–256. doi:10.1037/1528-3542.6.2.246

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, *194*, 465–475. doi:10.1007/s00221-009-1721-9

Smilek, D., Frischen, A., Reynolds, M. G., Gerritsen, C., & Eastwood, J. D. (2007). What influences visual search efficiency? Disentangling contributions of preattentive and post-attentive processes. *Perception and Psychophysics*, *69*, 1105–1116. doi:10.3758/BF03193948

Thornton, I. M., & Kourtzi, Z. (2002). A matching advantage for dynamic faces. *Perception*, *31*, 113–132. doi:10.1068/p3300

Tipples, J., Atkinson, A. P., & Young, A. W. (2002). The raised eyebrow: A salient social signal. *Emotion*, *2*, 288–296. doi:10.1037/1528-3542.2.3.288

Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249. doi:10.1016/j.psychres.2008.05.006

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, *1284*, 100–115. doi:10.1016/j.brainres.2009.05.075

Truong, K. P., van Leeuwen, D. A., & Neerincx, M. A. (2007). Unobtrusive multimodal emotion detection in adaptive interfaces: Speech and facial expressions 2. Short overview: State of the art. *Foundations of Augmented Cognition*, *4565*, 354–363.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*, 1395–1407. doi:10.1016/j.neunet.2006.10.001

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, *78*, 105–119. doi:10.1037/0022-3514.78.1.105

White, M. (1995). Preattentive analysis of facial expressions of emotion. *Cognition and Emotion*, *9*, 439–460. doi:10.1080/02699939508408975

Williams, M., Moss, S., Bradshaw, J., & Mattingley, J. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual Cognition*, *12*, 29–50.

Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception and Psychophysics*, *63*, 381–389.

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

Yoshikawa, S., & Sato, W. (2008). Dynamic facial expressions of emotion induce representational momentum. *Cognitive, Affective, and Behavioral Neuroscience*, *8*(1), 25–31.

APPENDIX A

Estimating motion in image sequences

The amount of movement was evaluated by comparing the first (neutral) and the last (full emotion) frame of the video sequence for each face identity and each expression. This comparison was carried out by computing the scalar quantity diff(A', B') defined as follows.

Let $A \in N^{n \times n}$ and $B \in N^{n \times n}$ be two square matrixes of size n, such that A_{ij} and B_{ij} represent the intensity of the i, j pixel for the greyscale images A and B, respectively. Having divided the A and B matrices into a grid of smaller blocks of size $m \times m$, with $m < n$, let us define the square density matrices $A'_{ij} \in Q^{\frac{n}{m} \times \frac{n}{m}}, B'_{ij} \in Q^{\frac{n}{m} \times \frac{n}{m}}$ such that:

$$A'_{ij} = \frac{\sum_{l=(i-1) \cdot m+1}^{i \cdot m} \sum_{k=(j-1) \cdot m+1}^{j \cdot m} A_{lk}}{\sum_{l,k=1}^{n} A_{lk}} \tag{A1}$$

and

$$B'_{ij} = \frac{\sum_{l=(i-1) \cdot m+1}^{i \cdot m} \sum_{k=(j-1) \cdot m+1}^{j \cdot m} B_{lk}}{\sum_{l,k=1}^{n} B_{lk}} \tag{A2}$$

represent the "density" of the images A and B within each single block of the grid. The sum of all the elements of each density matrix is unitary (i.e., $\sum_{ij} A'_{ij} = \sum_{ij} B'_{ij} = 1$). The distance between the images A and B is given by the sum of the absolute values of the difference between the corresponding density matrices:

$$diff(A', B') = \sum_{i=1}^{n/m} \sum_{j=1}^{n/m} |A_{ij} - B_{ij}| \tag{A3}$$

Given the inequality $0 \leq |a - b| \leq |a| + |b|$ it can be shown that:

$$0 \leq diff(A', B') \leq 2$$
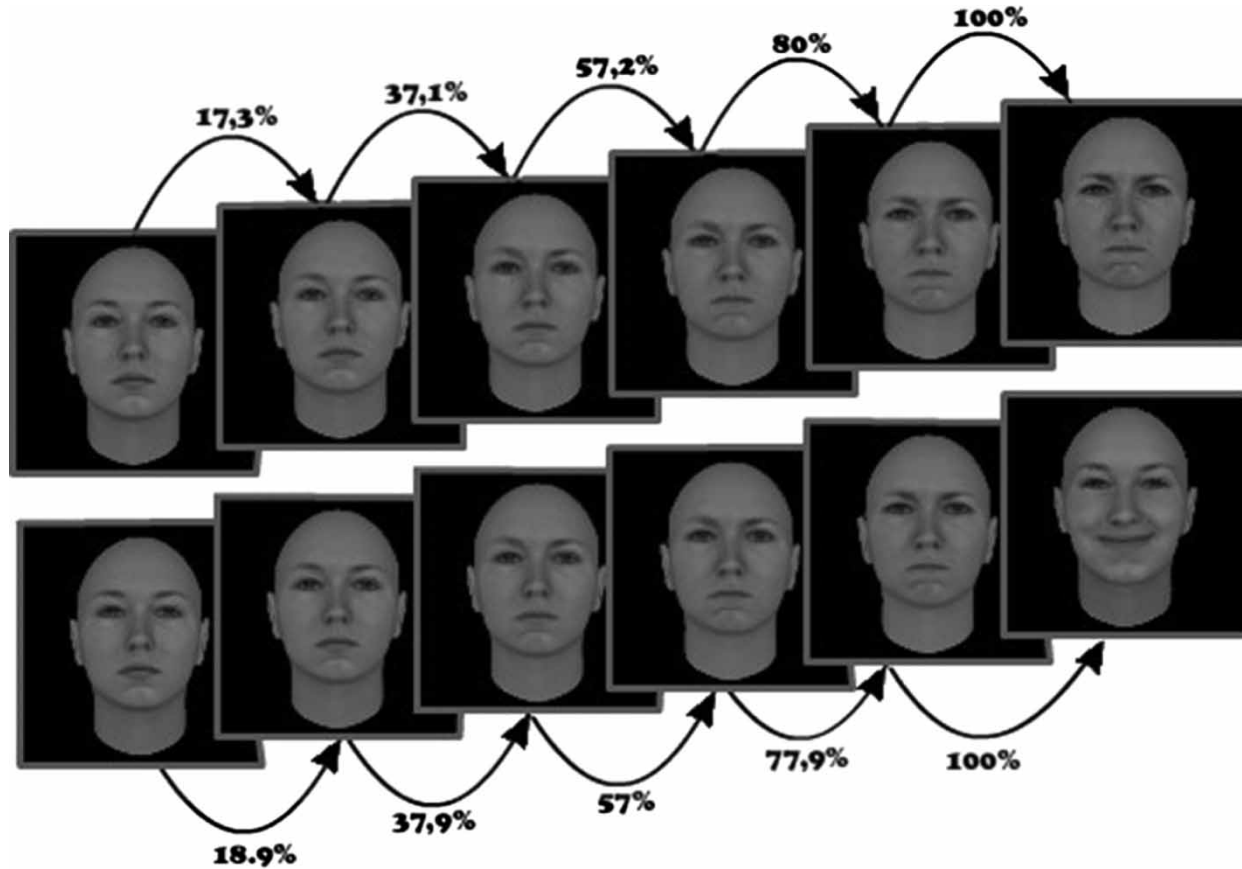
An example is shown in Figure A1.

Figure A1 Amount of movement computed according to Equation A3 for the transition between the neutral expression and the full angry expression (top) and the full happy expression (bottom) for one of the nine facial identities used in Experiments 1A, 1B, and 2. The numbers indicate the percentage of the total amount of movement between the frames 1–4, 4–7, 7–10, 10–13, 13–16 of the motion sequence. In the frame sequence represented at the bottom, the total amount of movement (from the neutral expression to the full-emotion expression) is 11% larger than in the frame sequence represented on the top.

APPENDIX B

Bottom-up saliency for schematic faces

The schematic faces used by Öhman et al. (2001) have been used in many investigations on the ASE (Calvo, Avero, & Lundqvist, 2006; Horstmann, 2007; Juth, Lundqvist, Karlsson, & Öhman, 2005; Lundqvist & Öhman, 2005; Mather & Knight, 2006; Tipples, Atkinson, & Young, 2002). By using iLab Neuromorphic Vision C++ Toolkit, Calvo and Nummenmaa (2008) determined that, within this set of schematic faces, the angry faces have a larger of bottom-up saliency than any of the other emotional expressions. The present analysis was run to determine whether the results of Calvo and Nummenmaa can be replicated by using the algorithm implemented in the SaliencyToolbox 2.2 with default settings.

We examined the six conditions of Experiment 2 of Öhman et al. (2001), in which a neutral, happy, or angry target face was shown in a crowd of distractors having a different emotional expression. As in the original experiment, the schematic faces were arranged in a 3×3 matrix. For each of the six condition (see Table B1), we created nine matrices by placing the target face in one of the nine cells of the matrix. We then computed the saliency map for the resulting 54 matrices (9 target positions $\times 6$ conditions). From each saliency map, we calculated the total activation in the target region (84×98 pixels). The average target activation for the six conditions is shown in Table B1. The results indicate that the bottom-up saliency of the target is 16.6% larger for an angry face in a neutral crowd than for a happy face in a neutral crowd. Moreover, the saliency of the target is 55.4% larger for an angry face in a happy crowd than for a happy face in an angry crowd. These results replicate those found by Calvo and Nummenmaa (2008).

TABLE B1
Bottom-up saliency for the schematic faces used by Öhman et al. (2001)

Condition	Mean saliency value
Angry face in a neutral crowd	12.67
Angry face in a happy crowd	10.19
Happy face in a angry crowd	5.67
Happy face in a neutral crowd	9.37
Neutral face in a angry crowd	5.62
Neutral face in a happy crowd	7.51