Journal of Experimental Psychology: General
2008, Vol. 137, No. 3, 471–494

0096-3445/08/$12.00 DOI: 10.1037/a0012771

Detection of Emotional Faces: Salient Physical Features Guide Effective Visual Search

Manuel G. Calvo
University of La Laguna

Lauri Nummenmaa
MRC Cognition and Brain Sciences Unit

In this study, the authors investigated how salient visual features capture attention and facilitate detection of emotional facial expressions. In a visual search task, a target emotional face (happy, disgusted, fearful, angry, sad, or surprised) was presented in an array of neutral faces. Faster detection of happy and, to a lesser extent, surprised and disgusted faces was found both under upright and inverted display conditions. Inversion slowed down the detection of these faces less than that of others (fearful, angry, and sad). Accordingly, the detection advantage involves processing of featural rather than configural information. The facial features responsible for the detection advantage are located in the mouth rather than the eye region. Computationally modeled visual saliency predicted both attentional orienting and detection. Saliency was greatest for the faces (happy) and regions (mouth) that were fixated earlier and detected faster, and there was close correspondence between the onset of the modeled saliency peak and the time at which observers initially fixated the faces. The authors conclude that visual saliency of specific facial features—especially the smiling mouth—is responsible for facilitated initial orienting, which thus shortens detection.

Keywords: facial expression, emotion, visual search, eye movements, saliency

A major function of selective attention is to prioritize the processing of important information at the expense of competing distractors. For adaptive reasons and because of their ubiquity, faces are probably the most biologically and socially significant visual stimuli for humans. Emotional expressions add further meaning to faces as they reveal the state, intentions, and needs of people and, therefore, indicate what observers can expect and how to adjust their own behavior accordingly. This makes emotional faces an ideal candidate for enhanced processing. Consistent with this view, neurophysiological research has found that emotional information from faces is detected rapidly 100 ms after stimulus onset, and different facial expressions are discriminated within an additional 100 ms (see reviews in Eimer & Holmes, 2007; Palermo & Rhodes, 2007). In the current study, we investigated *why* some emotional faces can be detected faster than others in a crowd and what properties of the facial expressions guide the search efficiently. A major issue is how detection is governed by a mechanism that is sensitive to salient visual features of some faces and facial regions and that subsequently triggers rapid shifts of attention to the salient features.

Manuel G. Calvo, Department of Cognitive Psychology, University of La Laguna, Tenerife, Spain; Lauri Nummenmaa, MRC (Medical Research Council) Cognition and Brain Sciences Unit, Cambridge, England.

This research was supported by Spanish Ministry of Education and Science Grant SEJ2004-420/PSIC to Manuel G. Calvo and Academy of Finland Grant 121031 to Lauri Nummenmaa. We thank Margaret Dowens and Gernot Horstmann for their helpful comments.

Correspondence concerning this article should be addressed to Manuel G. Calvo, Department of Cognitive Psychology, University of La Laguna, 38205 Tenerife, Spain. E-mail: mgcalvo@ull.es

An Advantage in the Detection of Some Emotional Faces

An initial step in the selective enhancement of stimulus processing involves fast detection of a target among distractors. The visual search paradigm has been used to investigate this process (see Müller & Krummenacher, 2006). With emotional face stimuli, this paradigm has produced mixed findings (for a review, see Frischen, Eastwood, & Smilek, in press). For *schematic* faces (i.e., line drawings) as stimuli, an angry face superiority has been typically found. Schematic angry (or negative-emotion) expressions are detected faster as discrepant targets among neutral expressions than vice versa, or in comparison with happy (or positive-emotion) targets (Calvo, Avero, & Lundqvist, 2006; Eastwood, Smilek, & Merikle, 2001; Fox et al., 2000; Horstmann, 2007; Juth, Lundqvist, Karlsson, & Öhman, 2005; Lundqvist & Öhman, 2005; Mather & Knight, 2006; Öhman, Lundqvist, & Esteves, 2001; Schubö, Gendolla, Meinecke, & Abele, 2006; Smilek, Frischen, Reynolds, Gerritsen, & Eastwood, 2007; Tipples, Atkinson, & Young, 2002). However, the external validity of schematic face stimuli is controversial (see Horstmann & Bauland, 2006). In fact, Juth et al. (2005) observed strikingly different effects for visual search of real versus schematic faces. With photographs of *real* faces, some studies have found an angry face advantage (Fox & Damjanovic, 2006; Hansen & Hansen, 1988; Horstmann & Bauland, 2006), although others have not (Purcell, Stewart, & Skov, 1996). Juth et al. obtained opposite results, that is, a happy face advantage, with discrepant happy expressions detected more quickly and accurately than angry and fearful targets in a context of neutral expressions. Similarly, Byrne and Eysenck (1995) reported a happy face superiority for a nonanxious group, with no differences between angry and happy faces for a high-anxious group. Gilboa-Schechtman, Foa, and Amir (1999) observed an angry face superiority over happy and disgusted faces for social-phobic participants but not for nonphobic controls. Williams, Moss, Bradshaw, and Mattingley (2005) found an advantage of both angry and happy faces (with no consistent difference between them) over sad and fearful faces.

From this review, we can conclude that both an angry and a happy face superiority has been observed in visual search tasks using photographic face stimuli. These findings also indicate that not all emotional expressions have been "created equal" in that some of them are detected faster and more accurately than others. There is, however, a limitation in this respect as no prior study has compared the detection of all six basic emotional facial expressions (fear, sadness, happiness, anger, disgust, and surprise; Ekman & Friesen, 1976). For schematic faces, angry and happy (and occasionally sad) expressions have typically been presented. For real faces, generally, angry and happy expressions have been used, except in Gilboa-Schechtman et al.'s (1999; happy, angry, and disgusted), Juth et al.'s (2005; happy, angry, and fearful), and Williams et al.'s (2005; happy, angry, sad, and fearful) studies. A related limitation is concerned with the fact that the face stimulus sample has usually been small and, thus, probably biased with respect to the representativeness of the natural variability of emotional expressions. For schematic faces, a single prototype of each expression was used in most of the studies. Regarding real faces, 12 or fewer (often, only two or three) different models have usually been presented (Byrne & Eysenck, 1995; Fox & Damjanovic, 2006; Gilboa-Schechtman et al., 1999; Hansen & Hansen, 1988; Horstmann & Bauland, 2006; Purcell et al., 1996; Williams et al., 2005). Only Juth et al. (2005) employed a large, 60-model sample. This limitation could probably account for the discrepancies regarding the superior detection of angry versus happy faces (see the General Discussion section). An approach that examines the search advantage of some expressions would thus benefit from comparing all six basic emotional expressions and from using a sufficiently varied and representative sample of stimuli.

Alternative Accounts for Visual Search Advantages in Emotional Face Processing

In this study, we investigate the factors and mechanisms responsible for the superior detection of some emotional expressions. A widely accepted view argues that the search advantage of certain expressions results from rapid processing of their affective significance. The framework to account for the findings of the angry face advantage was proposed by Öhman and collaborators (see Öhman & Mineka, 2001): Essentially, a *fear module* in the brain would preferentially process fear-relevant stimuli that have been phylogenetically associated with danger. Angry facial expressions are among these fear-relevant stimuli. They are detected quickly because they are signals of danger, and their prompt detection enables fast adaptive responses to avoid harm. Obviously, this argument cannot be applied directly to explain the happy face advantage. Nevertheless, such advantage would be instrumental in maximizing the receipt of social reward or establishing alliance and collaboration, thus quick detection of happy faces would also serve a general adaptive function. An important issue to be noted is that this explanation—as applied to either angry or happy faces—implies that *emotional meaning* is responsible for visual search advantages (see Reynolds, Eastwood, Partanen, Frischen, & Smilek, in press). In line with this, Lundqvist and Öhman (2005) have argued that the correlation between search performance and valence ratings of schematic faces is consistent with the hypothesis that the affective significance of the faces underlies the detection superiority. Similarly, the priming of probe *words* by semantically congruent schematic *faces* suggests that the enhanced detection of unambiguous emotional faces involves processing of the meaning of the expressions, not merely discrimination of formal visual features (Calvo & Esteves, 2005).

There is, however, an alternative view arguing that visual search of faces is not guided by the processing of affective meaning. Instead, the efficient search of certain expressions could be accounted for by perceptual rather than affective factors. The visual search task involves detection of a discrepant target among distractor stimuli, and *visual* discriminability between the target and the distractors is a major determinant of performance (Duncan & Humphreys, 1989). Discriminability could determine visual search differences between facial expressions at three levels: purely visual saliency, featural, and configural. The three alternative accounts are complementary, rather than mutually exclusive, in so much as they involve visual processing of facial stimuli at different levels of increasing perceptual complexity. Nevertheless, whereas the configural conceptualization—and, to a lesser extent, the featural notion—could accommodate the encoding of meaning of the emotional expressions, this would be incompatible with the purely perceptual saliency explanation.

First, according to a *visual saliency* account, the discriminability at early stages of visual processing arises from the physical saliency of the target (Nothdurft, 2006). The stimulus properties that guide the initial stages of search are those that can be rapidly detected by the primary visual (V1) cortex (e.g., luminance, color, and orientation; Itti & Koch, 2000). Importantly, none of these low-level stimulus properties is meaningful in a strict sense, and they are thus devoid of any emotional significance. The processing of such properties proceeds in a bottom-up rather than a top-down fashion. When applied to face visual search, this approach implies that the search advantage could be due to a greater visual salience of a particular target emotional expression than others, in a context of neutral faces.

Second, according to a *featural* account, the search advantage of certain emotional expressions can be due to their better discriminability from the neutral distractors at the level of single facial areas or features, such as upturned lip corners, open eyes, or frowning. Facial features represent particular combinations of low-level image properties that produce specific shapes and, thus, constitute significant units or components of the faces; however, the representation of these features is encoded later in the ventral visual stream (McCarthy, Puce, Belger, & Allison, 1999). These features could be particularly prominent or distinctive in some emotional expressions when presented in an array of neutral expressions. These single features might have acquired some affective meaning through association with the whole facial expression in which they typically appear (see Cave & Batty, 2006). However, they can probably be readily detected regardless of meaning, only on the basis of physical differences with respect to the corresponding neutral feature (e.g., closed lips) of the other faces in the context.

Finally, according to a *configural* account, discriminability could involve the facial configuration, that is, the whole facial expression. *Configural information* refers to the structural relationship between different facial features (e.g., the relative shape and positioning of the mouth in relation to those of the nose, eyes, etc.; Carey & Diamond, 1977). This is the spatial information that makes a face a face. The identification of facial expressions is based mainly on configural processing, although featural processing also plays a role (Calder, Young, Keane, & Dean, 2000). An important question is, however, the extent to which visual search can be performed on the basis of simple *detection* of a discrepant facial configuration (i.e., that a target face is different) without the need for *identification* (i.e., what kind of expression it is; see Lewis & Edmonds, 2005). Certain facial configurations may just be more visually distinctive than others, and this could facilitate detection without expression encoding.

The Current Study: The Roles of Visual, Featural, and Configural Factors

We conducted a series of seven experiments to distinguish between the roles of these three levels of face processing in visual search. Specifically, we set out to determine which properties of the different emotional faces can guide search and facilitate detection. To examine the role of low-level visual discriminability, we compared emotional and the respective neutral facial expressions on five physical image characteristics (luminance, contrast, global energy, color, and texture; Experiment 1), and we explored the effect of a white versus black background display (Experiment 7). In a more elaborate approach, we combined some of the image characteristics into an overall or "master" saliency map of the whole visual array of faces (Experiment 2) or of different regions within each face (Experiment 6). Low-level image properties and saliency have been found to influence the initial covert and overt shifts of visual attention while inspecting pictorial stimuli (Itti & Koch, 2000; Parkhurst, Law, & Niebur, 2002). Accordingly, such image properties are also expected to influence visual search and detection. If the advantage of angry or happy (or any other) facial expression is due to low-level discriminability, differences in physical properties and saliency between the angry or the happy targets and the corresponding neutral faces should be greater than for other emotional faces. To our knowledge, no prior study has addressed the issue of whether and how perceptual salience can be responsible for the search and detection of emotional faces.

To investigate the role of featural versus configural processing, we employed two different methods. In the first approach, inverted (i.e., upside-down) face arrays were presented for visual search and were compared with upright arrays (Experiment 3). We assumed that inversion preserves low-level visual properties and has minimal impact on the processing of single features but that it dramatically impairs configural processing and facial expression recognition (Farah, Tanaka, & Drain, 1995). In contrast, upright presentation preserves the facial configuration (in addition to low-level properties and features) and therefore allows for holistic or configural processing of the expression. If the search advantage of angry and happy (or any other) facial expressions is due to featural processing, inversion will be less detrimental for the search of angry or happy faces than for the other faces. In contrast, if the advantage involves configural processing, performance will be particularly impaired by inversion. Inverted versus upright paradigms have been used in prior research with real face stimuli in visual search tasks (Fox & Damjanovic, 2006; Horstmann & Bauland, 2006). The results, however, have been equivocal, with inversion either eliminating (Fox & Damjanovic, 2006) or not eliminating (Horstmann & Bauland, 2006) the superiority of angry over happy faces.

In the second approach, we explored the roles of the eyes and the mouth in visual search performance. These regions were presented alone (Experiments 4A and 4B), and detection performance for them was compared with that for the whole face (Experiment 1). If the facilitated search of some expressions is contingent on configural processing, the search advantage will occur only when the whole face is presented. If, in contrast, the effect is based on the processing of single facial components, the presentation of single regions of the faces will produce a similar advantage to the whole face. This approach has also been used in prior studies with real face stimuli (Fox & Damjanovic, 2006; Horstmann & Bauland, 2006), although the findings have been inconsistent: The eye, but not the mouth region (Fox & Damjanovic, 2006), or the mouth, but not the eye region (Horstmann & Bauland, 2006), has been reported to produce an angry face superiority effect. In an attempt to clarify and extend the role of significant parts of the faces, we also used a variant of the procedure, which involved *removing* the eye or the mouth regions from the whole face (Experiments 5A and 5B). If a face region is necessary for producing a search advantage, removing such region will eliminate the advantage of an emotional expression.

Experiment 1

Emotional Face Detection: Searching for a Detection Advantage

In Experiment 1, we investigated whether visual search performance varies as a function of emotional facial expression. This experiment served to establish the basic paradigm and also the basic findings (i.e., whether some expressions are detected faster than others) for which alternative explanations were examined in the following experiments. To expand the comparisons beyond previous research, we used all six basic emotional expressions. To increase generalizability, we used a large stimulus sample of 24 different posers. Visual arrays composed of one emotional target face and six neutral context faces (or all seven neutral faces) were presented for target detection.

Method

Participants. Twenty-four psychology undergraduates (18 women, 6 men; from 19 to 23 years of age) participated for course credit. The participants for Experiments 1–7 were recruited from the University of La Laguna (Tenerife, Spain).

Stimuli. The stimuli were 168 digitized color photographs selected from the Karolinska Directed Emotional Faces (KDEF; Lundqvist, Flykt, & Öhman, 1998; see http://www.facialstimuli.com/). A sample of the pictures is shown in Figure 1. The stimuli portrayed 24 different individuals (12 women, 12 men) each posing seven expressions (neutral, happy, angry, sad, disgusted, surprised, and fearful) gazing directly at the viewer. Four additional individuals (2 women, 2 men; 28 photographs) were used for practice trials. These models were amateur actors with a mean age of 25 years (range = 20–30 years) and of Caucasian origin. According to the authors of the KDEF (Lundqvist et al., 1998), all

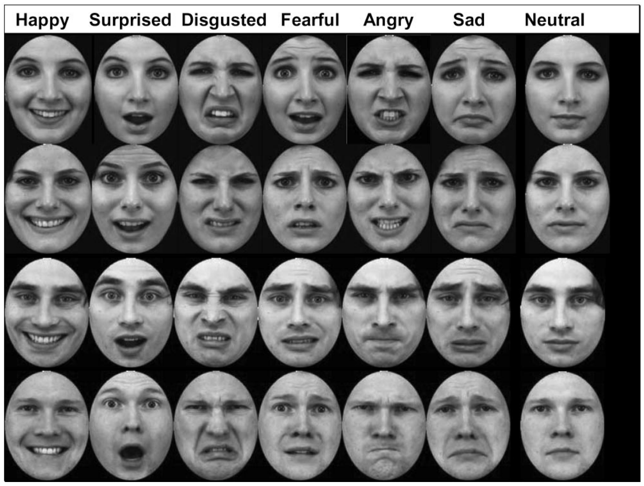

Figure 1. Sample Karolinska Directed Emotional Faces pictures used in the current study.

the models received written instructions entailing a description of the seven expressions and were asked to rehearse these for 1 hr before coming to the photo session. It was emphasized that they should try to evoke the emotion that was to be expressed and—while maintaining a way of expressing the emotion that felt natural to them—to try to make the expression strong and clear. The 24 selected models were those who proved to best convey the different emotional expressions in a previous recognition study (Calvo & Lundqvist, 2008; recognition rates ranged between 80% and 97%). We used the following KDEF pictures for the experimental trials—women: 01, 02, 03, 05, 07, 09, 13, 14, 19, 20, 29, 31; men: 03, 05, 08, 10, 11, 12, 14, 17, 23, 29, 31, 34.

Each photograph was cropped: Nonfacial areas (e.g., hair, neck, etc.) were removed by applying an ellipsoidal mask (see Williams et al., 2005). Stimulus displays were arranged in a circle such that each array contained six faces surrounding a central face of the same model (see Figure 2). Each face subtended a visual angle of 3.8° × 3.0° at a 60-cm viewing distance. The center of the central face coincided spatially with the starting fixation point. The center of all the surrounding faces was located at 3.8° from this fixation point and from the two adjacent faces. The faces appeared against a black background.

There were two types of stimulus displays. The display of specific interest involved one discrepant emotional target face among six neutral faces. For these trials, the central face was always neutral, and the emotional target appeared in one of the six surrounding locations. Each participant was presented with 144 trials of this kind, with one face of each emotional expression of each model. Target location was counterbalanced. In an additional type of array (72 trials), all seven faces were neutral, with the same model presented on three occasions. Trials were randomly assigned to three blocks and randomly presented within each block.

Design, procedure, and measures. There were two within-subjects factors for displays with one discrepant target: expression of the target face (happy vs. angry vs. sad vs. disgusted vs. surprised vs. fearful) and target location in the array (left vs. middle vs. right). Each target appeared once in each location for each participant. To explore potential lateralization effects, we averaged scores for the two leftwards locations, the two rightwards locations, and the central upwards and downwards vertical locations (see Williams et al., 2005).

The stimuli were presented on a 17-in. (43.18-cm) Super video graphics array (VGA) monitor, connected to a Pentium-IV 3.2-GHz computer. Stimulus presentation and data collection were controlled by the E-Prime experimental software (Schneider, Eschman, & Zuccolotto, 2002). Each trial (see Figure 2) started with a central fixation cross for 500 ms. Following offset of the cross, the face display appeared and remained until the participant responded. The task involved pressing one of two keys to indicate whether there was a discrepant face in the array. Visual search performance was assessed by response accuracy and reaction times from the onset of the stimulus display until the participant's response.

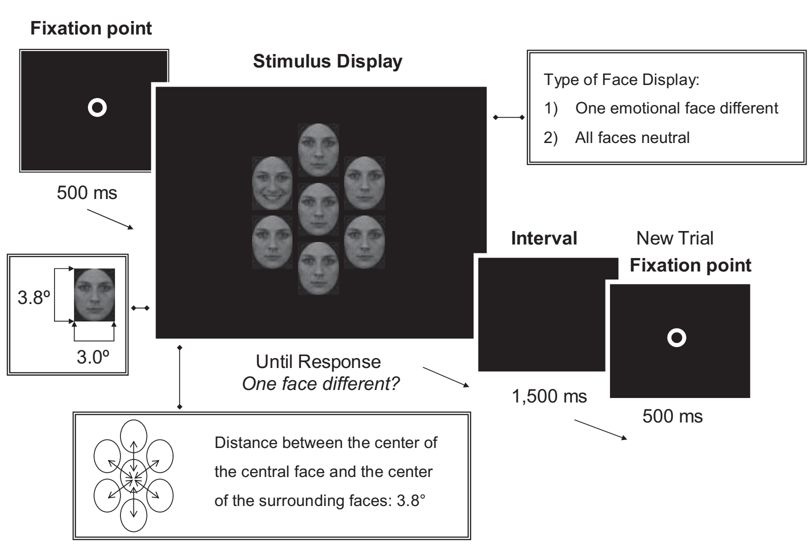

Figure 2. Sequence of events and overview of basic characteristics of a trial.

Assessment of low-level image properties. We compared each emotional face and the corresponding neutral face on several physical properties, to examine the possibility that some emotional target faces might differ more than others from the neutral context faces, and that this could account for the visual search advantages. We computed basic image statistics, such as mean luminance, contrast density (root-mean-square contrast), and global energy (see Kirchner & Thorpe, 2006) with Matlab 7.0 (The Mathworks, Natick, MA). In addition, we computed color and texture similarity with local pixel-by-pixel principal component analysis with reversible illumination normalization (see Latecki, Rajagopal, & Gross, 2005).

Results

Response accuracy and detection times for correct responses. The probability of correct responses and the search times were analyzed by means of 6 (emotional expression of target) $\times$ 3 (target location) analyses of variance (ANOVAs). Bonferroni corrections and alpha levels of $p < .05$ were used for all multiple contrasts in this and all the following experiments. Mean scores and statistical significance of the contrasts (indicated by superscripts) are shown in Table 1.

For *response accuracy*, there was a facial expression effect, $F(5, 115) = 20.71$, $p < .0001$, $\eta_p^2 = .47$, with no spatial location effect or an interaction (Fs ≤ 1; henceforth, only statistically significant effects are reported). As indicated in Table 1, accuracy was highest for happy and surprised targets, followed by disgusted and fearful targets, and poorest for angry and sad targets. For *response times*, a significant effect of expression, $F(5, 115) = 52.17$, $p < .0001$, $\eta_p^2 = .69$, emerged. As indicated in Table 1, responses were fastest for happy targets, followed by surprised, disgusted, and fearful targets, which were faster than for angry targets and were slowest for sad targets.

Analysis of low-level image properties. Differences in luminance, root-mean-square contrast, energy, color, and texture were computed between the neutral face and each of the emotional faces of the same model. Mean scores are presented in Table 2. One-way ANOVAs (6: emotional expression) were conducted on these difference scores. For luminance, the effect did not reach statistical significance, $F(5, 115) = 2.17$, $p = .078$, with only a tendency for the happy faces to be more similar to the neutral faces than were the other emotional faces. For contrast density, no differences emerged (all multiple contrasts, $ps \geq .11$). For energy, a significant effect, $F(5, 115) = 8.53$, $p < .0001$, $\eta_p^2 = .27$, indicated that the surprised and the happy faces were more similar to the neutral faces than were the other emotional faces. For color and texture, no significant differences appeared between the different expressions (color: $p = .32$; texture: $p = .19$).

Discussion

There were significant differences in visual search performance among most of the emotional faces. Regarding the two most investigated expressions, that is, happy and angry, the results were clear-cut. Target faces with happy expressions were responded to faster than any other target and also with greater accuracy. The advantage of happy, relative to angry, faces is consistent with findings in some prior studies (Byrne & Eysenck, 1995; Calvo, Nummenmaa, & Avero, in press; Juth et al., 2005) but is in contrast to others showing an anger superiority both for real faces (Fox & Damjanovic, 2006; Hansen & Hansen, 1988; Horstmann & Bauland, 2006) and schematic faces (e.g., Calvo et al., 2006; Lundqvist & Öhman, 2005). In the General Discussion section, we address the explanation of these discrepancies, once we have examined additional evidence.

Beyond the two most investigated expressions, the present findings extend the comparison to six expressions. In addition to the happy faces, the surprised and, to a lesser extent, the disgusted faces were detected better than the fearful and the angry faces, whereas the sad faces were detected most poorly. These detection differences were not directly related to differences in low-level image properties. It is possible, however, that each image property in isolation does not account for detection because the visual system may *combine* the measured properties in a nonlinear fashion—or global image statistics for the whole face may not be sensitive to local differences between facial *regions*, which may, nevertheless, be perceptually salient. We examined these possibilities in Experiments 2 and 6, respectively, by using a computationally modeled visual saliency that combines several image properties.

Table 1
Mean Probability of Correct Responses and Reaction Times in the Visual Search Task, as a Function of Type of Emotional Expression of the Target Face, in Experiment 1

	Type of expression					
Variable	Happy	Surprised	Disgusted	Fearful	Angry	Sad
Accuracy (probability)						
M	.981 [a]	.977 [a]	.962 [ab]	.932 [b]	.885 [c]	.867 [c]
SD	.037	.027	.061	.072	.084	.108
Response times (in milliseconds)						
M	741 [a]	816 [b]	827 [bc]	886 [c]	959 [d]	1,082 [e]
SD	142	214	171	204	220	214

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent. Bonferroni corrections ($p < .05$) were used for all multiple contrasts and experiments.

Table 2
Mean Luminance, RMS Contrast, Energy, Color, and Texture Difference Scores Between Neutral and each Type of Emotional Face Stimulus (i.e., Neutral − Emotional) for the Whole-Face Stimuli Used in Experiment 1

	Type of expression					
Variable	Happy	Surprised	Disgusted	Fearful	Angry	Sad
Luminance	4.74	6.71	6.79	6.13	6.05	6.12
RMS contrast	0.015	0.023	0.019	0.021	0.011	0.012
Energy ($\times 10^{-5}$)	50^{a}	45^{a}	149^{b}	124^{b}	149^{b}	154^{b}
Color	167	196	208	196	202	194
Texture	−0.443	0.136	0.241	−0.049	0.172	−0.059

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent. RMS = root-mean-square.

Experiment 2

Visual Saliency of Emotional Faces and Attentional Orienting

In Experiment 2, we investigated whether visual saliency accounts for the search advantage of some emotional expressions and whether this occurs through an effect on initial orienting or later decision processes. To this end, we first computed a global saliency map of each face array with the algorithm developed by Itti and Koch (2000; see also Itti, 2006; Navalpakkam & Itti, 2005). The saliency map represents the relative visual conspicuity of the different parts of the image. Second, we used eye-movement monitoring during the search task to distinguish between an orienting stage (from onset of the face array until the first fixation on the target) and a decision stage (from first fixation on the target until the manual response). Various models have proposed that saliency influences initial shifts of covert and overt attention (see Torralba, Oliva, Castelhano, & Henderson, 2006). Supporting evidence has shown that the initial distribution of eye fixations on a picture is determined by the saliency weights of the different parts of the image (e.g., Parkhurst et al., 2002; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006). Accordingly, if visual saliency is responsible for the search advantage of happy faces, then (a) happy face targets will have greater saliency values than any other emotional target in an array of neutral faces, and (b) happy targets will receive the first eye fixation more often and earlier than the other targets, which would thus shorten the whole detection process.

Method

Participants. Twenty-four undergraduate psychology students (18 women, 6 men; from 19 to 22 years of age; 21 right handed) participated for course credit.

Stimuli. In addition to the 24 KDEF models used in Experiment 1, four more models were included (women: no. 11 and no. 26; men: no. 06 and no. 13), each posing a neutral and the six emotional expressions.

Apparatus, procedure, and design. The stimuli were presented on a 21-in. (53.34-cm) monitor with a 120-Hz refresh rate, connected to a Pentium IV 3.2-GHz computer. Participants' eye movements were recorded with an EyeLinkII tracker (SR Research Ltd., Mississauga, Ontario, Canada), connected to a Pentium IV 2.8-GHz computer. The sampling rate of the eyetracker was 500 Hz, and the spatial accuracy was better than 0.5°, with a 0.01° resolution in pupil tracking mode. A forehead and a chin rest were used to keep viewing distance constant (60 cm). Each trial started with a central drift correction circle (0.8° of diameter). When the participant fixated this circle, the face display appeared and remained until the participant's response. The procedure and the design were otherwise identical to those in Experiment 1.

Measures. In addition to response accuracy and reaction times, eye movement recordings were employed to construct three variables. At an early stage, attentional orienting was assessed by means of the *probability of first fixation*, that is, the likelihood that the initial fixation on the array landed on the discrepant target face, and *localization time,* that is, the time from the onset of the stimulus display until the target face was fixated. At a later stage, *decision time* was operationalized as the time from the first fixation on the target until the response was made. Thus, total performance time was decomposed into one process guiding visual search for the target (until fixation) and another involving detection of the target as different from the context (after fixation).

Visual saliency. We computed a purely bottom-up saliency map for each array of one discrepant emotional target and six neutral distractors by using the iLab Neuromorphic Vision C++ Toolkit (iNVT; see Itti & Koch, 2000). This neuromorphic model simulates which elements (and in which order) in a given scene attract the attention of human observers. Briefly, the visual input is first decomposed and processed by feature (e.g., local contrast, orientation, energy) detectors mimicking the response properties of retinal neurons, lateral geniculate nucleus, thalamus, and V1. These features are integrated for a neural saliency map that is a graded representation of the visual conspicuity of each pixel in the image. Salient areas (or objects) thus stand out from the background, including other surrounding objects. A winner-takes-all (WTA) neural network determines the point of highest saliency and draws the focus of attention to this target. The time taken for this depends on the saliency distribution in the neural saliency map, which is encoded by the WTA network: The more unambiguous the map, the faster the winning location is determined. To allow attention to shift to the next most salient target, an inhibition of return (IOR) is triggered for the currently attended object, reducing its saliency weight and resulting in a modified saliency map. The interplay between the WTA and IOR ensures that the saliency map is scanned in order of decreasing saliency, thus simulating how the allocation of attention would change. As applied to our stimulus arrays, we computed the initial (i.e., before the first IOR) saliency map as well as those before the second, third, fourth, and fifth IORs. The average modeled onset times for the respective IORs were 96, 161, 222, 311, and 558 ms.

Results

Visual search performance. The dependent variables were analyzed by means of 6 (emotional expression of target) × 3 (spatial location) ANOVAs. See mean scores and multiple contrasts in Table 3 and Figure 3.

An expression effect on *response accuracy*, $F(5, 115) = 40.99$, $p < .0001$, $\eta_p^2 = .64$, revealed more accurate detection of happy, surprised, disgusted, and fearful targets than of angry targets, all of which were detected more accurately than sad targets. Effects of expression on *response times*, $F(5, 115) = 63.84$, $p < .0001$, $\eta_p^2 = .74$, showed faster responses for happy targets, followed by surprised and disgusted targets, which were faster than for fearful and angry targets and were slowest for sad targets.

For *probability of first fixation*, main effects of expression, $F(5, 115) = 33.40$, $p < .0001$, $\eta_p^2 = .59$, indicated that happy targets were the most likely to be fixated first, followed by surprised and disgusted targets, all of which were more likely to be fixated first than fearful, angry, and sad targets. There was a significant negative correlation between probability of first fixation and response times, $r(144) = -.60$, $p < .0001$. For target *localization time*, effects of expression, $F(5, 115) = 44.63$, $p < .0001$, $\eta_p^2 = .66$, indicated that the time prior to fixation on the target was shortest for happy faces; it was shorter for surprised and disgusted faces than for fearful and angry faces, and it was longest for sad faces. There was a significant correlation between localization time and response times, $r(144) = .67$, $p < .0001$. For *decision time*, the effects of expression were still statistically significant, though considerably reduced, $F(5, 115) = 4.85$, $p < .01$, $\eta_p^2 = .17$. The only difference as a function of emotional expression involved slower decision responses for sad faces than for happy, surprised, and disgusted faces.

Target saliency. Mean target saliency scores across five time IORs from the onset of the display (see Figure 4) were analyzed by a 6 (expression) × 5 (IOR: 1 to 5) ANOVA. Main effects of expression, $F(5, 135) = 2.92$, $p < .05$, $\eta_p^2 = .10$, and IOR, $F(4, 108) = 4.69$, $p < .025$, $\eta_p^2 = .15$, were qualified by an interaction, $F(20, 540) = 3.33$, $p < .025$, $\eta_p^2 = .11$. To decompose the interaction, we conducted one-way (expression) ANOVAs for each IOR. No significant effects appeared for the first, second, third, and fifth IORs (all Fs < 1.85, $p > .16$). In contrast, for the fourth IOR (at 311 ms), a reliable effect of expression emerged, $F(5, 115) = 6.74$, $p < .01$, $\eta_p^2 = .20$. Multiple contrasts revealed significant differences between happy faces and all the other emotional faces ($ps < .05$; except the disgusted faces, $p = .15$).

Discussion

There were three major new findings. First, target faces with a happy expression were more often fixated first and localized earlier than any other emotional targets. Second, emotional expression affected search times at the initial stage involving target localization (i.e., first fixation) rather than at the later stage involving decision that the target was different from the context faces (i.e., response latency after first fixation). Third, happy targets were also more visually salient than the other target faces. It is interesting to note that the more likely a target was fixated first and the more quickly it was localized, the faster it was detected as different from the distractors. This reveals how much the initial orienting of attention to targets contributes to the final detection time. Furthermore, this finding is important if we want to determine how saliency can account for the search advantage of happy faces. Our findings suggest that the enhanced saliency of happy targets is directly responsible for the biased attentional orienting (i.e., first fixation and faster localization) to these faces. Because of this facilitation of orienting, saliency would then contribute to shorten the detection process, hence the advantage in response times.

The analysis of saliency scores across five periods clearly illustrates how saliency affects orienting. The saliency peak for the happy targets—and the significant saliency differences from the other targets—occurred at 311 ms (fourth IOR) following the array onset. This corresponds very closely to the participants' mean localization time, as the average time from stimulus onset to the first fixation on happy targets was 348 ms. Just slightly (i.e., 37 ms) after the modeled saliency arose in the location of the happy face, an eye fixation landed on the target face. Presumably, in-

Table 3
Mean Probability of Correct Responses, Total Correct Response Times, and Probability of First Fixation on the Target Face, as a Function of Type of Emotional Expression, in Experiment 2

	Type of expression					
Variable	Happy	Surprised	Disgusted	Fearful	Angry	Sad
Accuracy (probability)						
M	.969 [a]	.944 [a]	.932 [a]	.922 [a]	.819 [c]	.734 [d]
SD	.055	.081	.130	.128	.164	.166
Response times (in milliseconds)						
M	794 [a]	820 [ab]	838 [b]	940 [c]	971 [c]	1,077 [d]
SD	176	178	184	193	146	177
First fixation (probability)						
M	.625 [a]	.540 [b]	.535 [b]	.413 [c]	.337 [cd]	.297 [d]
SD	.178	.156	.178	.145	.123	.140

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent.

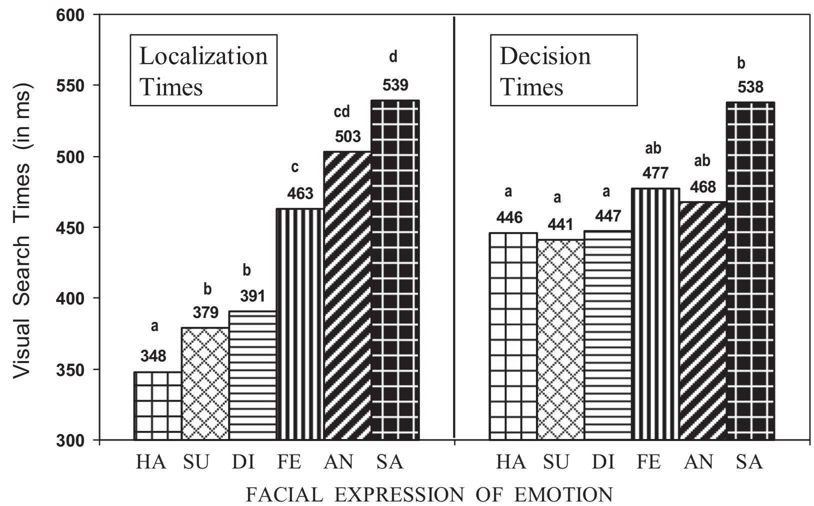

Figure 3. Target localization times and decision times as a function of emotional face expression in Experiment 2. Significant differences in multiple contrasts are indicated by superscripts. HA = happy; SU = surprised; DI = disgusted; FE = fearful; AN = angry; SA = sad.

creased saliency caused a shift in *covert* attention toward the target location and *subsequently* drove *overt* attention to the face. This is entirely consistent with the assumption that visual saliency is the main factor responsible for early orienting (Itti & Koch, 2000; Henderson, Weeks, & Hollingworth, 1999) and with data indicating that the most salient objects in a scene attract the initial eye fixations (Parkhurst et al., 2002; Underwood et al., 2006).

In sum, the current experiment has shown that the detection advantage of happy faces can be explained in terms of their higher visual saliency. This is consistent with the low-level account that we put forward in the introduction. We also proposed a featural and a configural account that could explain the search advantage of salient faces. To examine the extent to which the visual saliency explanation is compatible with the other two accounts, we conducted the following experiments.

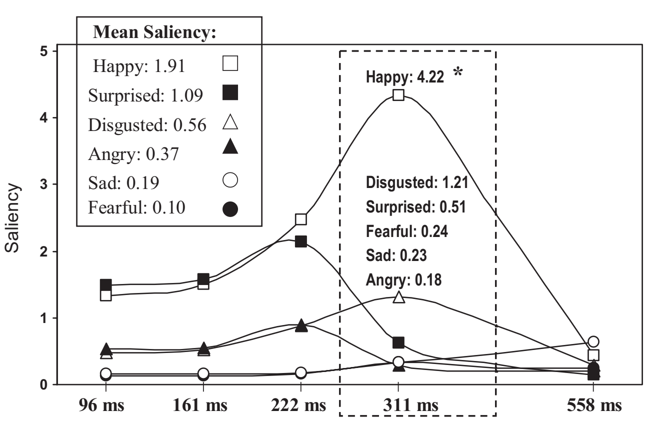

Figure 4. Mean saliency across five inhibition of returns (IORs) from onset of face display and scores at the fourth IOR (at 311 ms) as a function of emotional expression in Experiment 2.

Experiment 3

Inverted Versus Upright Faces: Featural Versus Configural Processing

Face identification is highly dependent on configural processing (Calder et al., 2000). Relative to upright faces, recognizing spatially inverted faces is surprisingly poor, with the impairment being larger for faces than for other stimuli (see Maurer, LeGrand, & Mondloch, 2002). It is assumed that inversion disrupts the holistic configuration but preserves the local facial features, and that an inverted face requires piecemeal processing of isolated features (Leder & Bruce, 1998). Accordingly, if the search advantage of happy (or other) facial expressions involves configural processing, the advantage will occur only when faces are presented upright. In contrast, if the advantage remains for inverted faces, some local features rather than emotional expression per se might be producing the effect. We addressed these hypotheses by presenting arrays of faces in a natural upright or in an inverted orientation.

Method

Participants. Forty-eight psychology students (34 women, 14 men; from 18 to 22 years of age) were randomly assigned to upright or inverted display conditions (24 participants each).

Stimuli, procedure, and design. The same KDEF photographs of individual faces as in Experiment 1 were presented. In the inverted condition, the arrays of faces were displayed upside-down. The design involved a between-subjects factor (orientation: upright vs. inverted) in addition to the two within-subjects factors (emotional expression and location of the target face). The procedure was otherwise the same as in Experiment 1.

Results

Trials with all faces identical. When all faces conveyed a neutral expression, responses were faster in the upright condition

(1,181 ms) than in the inverted condition (1,473 ms), $t(46) = 3.09$, $p < .01$, with no differences in response accuracy (.935 vs. .934, respectively).

Trials with one discrepant emotional target. The dependent variables were analyzed with 6 (emotional expression) × 3 (location) × 2 (orientation) mixed ANOVAs. For *response accuracy*, there were effects of expression, $F(5, 230) = 78.60$, $p < .0001$, $\eta_p^2 = .63$; orientation (mean probability of correct responses: upright = .896; inverted = .854), $F(1, 46) = 4.60$, $p < .05$, $\eta_p^2 = .091$; and an Expression × Orientation interaction, $F(5, 230) = 5.68$, $p < .0001$, $\eta_p^2 = .11$. To decompose the interaction, separate analyses for the upright and the inverted condition revealed strong effects of expression in both cases, $F(5, 115) = 33.21$, $p < .0001$, $\eta_p^2 = .59$, and, $F(5, 115) = 46.32$, $p < .0001$, $\eta_p^2 = .67$, respectively. As indicated in Figure 5, generally, accuracy was higher for happy targets, followed by surprised, disgusted, and fearful targets, than for angry targets, and it was poorest for sad targets. The interaction was due to the fact that in comparison with the upright orientation, inversion impaired detection of sad, $t(46) = 3.12$, $p < .01$, and angry, $t(46) = 2.47$, $p < .025$, expressions, but it did not affect the other expressions.

For *response times*, main effects of expression, $F(5, 230) = 138.42$, $p < .0001$, $\eta_p^2 = .75$, and orientation (upright: 909 ms; inverted: 1,039), $F(1, 46) = 5.58$, $p < .025$, $\eta_p^2 = .11$, were qualified by their interaction, $F(5, 230) = 9.92$, $p < .0001$, $\eta_p^2 = .18$. Separate analyses for the upright and the inverted condition revealed strong effects of expression in both cases, $F(5, 115) = 76.02$, $p < .0001$, $\eta_p^2 = .77$, and, $F(5, 115) = 73.56$, $p < .0001$, $\eta_p^2 = .76$, respectively. As indicated in Figure 5, the pattern of differences as a function of emotional expression was similar in the upright and the inverted condition. Response times were shorter for happy than for disgusted and surprised targets, which were shorter than for fearful and angry targets, and were longest for sad targets. The interaction was due to the fact that, in comparison with the upright orientation, inversion slowed detection of fearful, $t(46) = 2.02$, $p < .05$; angry, $t(46) = 2.98$, $p < .01$; and sad, $t(46) = 3.58$, $p < .001$, expressions, but it did not significantly affect the other expressions.

Discussion

Two interesting new findings emerged. First, the pattern of search differences between emotional expressions remained essentially the same in the inverted and the upright conditions. This is consistent with a featural rather than a configural explanation. Prior research comparing happy and angry expressions has provided mixed findings. Performance differences between faces have been found to remain under spatial inversion (with real faces as stimuli: Horstmann & Bauland, 2006; and with schematic faces: Horstmann, 2007; Öhman et al., 2001), which is therefore consistent with our own findings. In contrast, however, inversion has also been found to abolish those differences (with real faces: Fox & Damjanovic, 2006; with schematic faces: Eastwood, Smilek, & Merikle, 2003; Fox et al., 2000), which would support a configural explanation.

Second, importantly, not all emotional expressions were equally affected by inversion. Inversion delayed detection of fearful, angry, and sad targets, and it even resulted in impaired accuracy for sad and angry targets, in comparison with the upright condition. In contrast, inversion did not influence accuracy and response times for happy, surprised, and disgusted targets. A main conclusion is that visual search is guided by featural information of happy, surprised, and disgusted expressions to a greater extent than for fearful, angry, and, especially, sad expressions, which rely more on configural processing. This suggests that the happy (and surprised, and disgusted) face detection advantage is strongly dependent on the perception of single features rather than on emotional meaning.

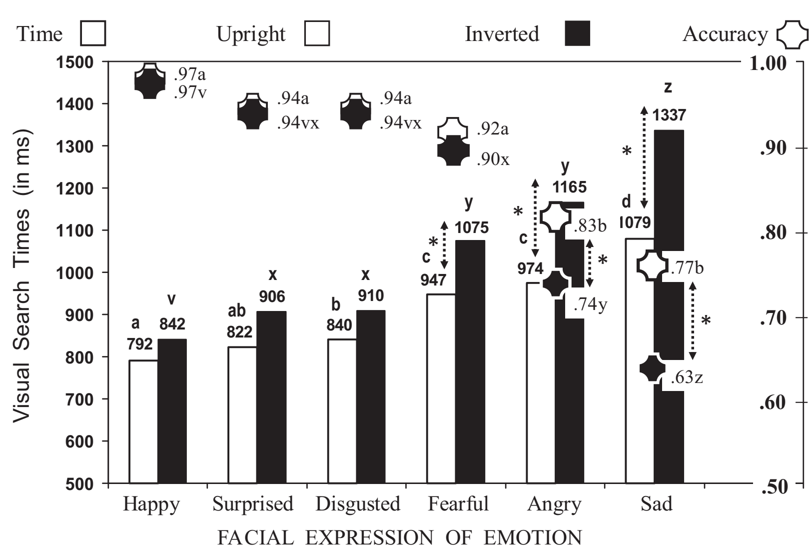

Figure 5. Mean response times and probability of correct responses as a function of emotional expression in Experiment 3. Significant differences in multiple contrasts are indicated by subscripts (a–d: upright condition; v–z: inverted condition). Vertical arrows and asterisks indicate significant differences between the upright and the inverted conditions.

To identify which facial features might be responsible for this advantage, we conducted Experiments 4 and 5.

Experiments 4A and 4B

Role of the Eyes and the Mouth: Sufficiency Criterion

In Experiments 4A–4B and 5A–5B, we examined the role of facial features in visual search by either presenting the eye or the mouth region alone or removing them from the face. We then compared these conditions with a whole-face display condition. Given that the participants belonged to the same pool as those in Experiment 1 and were randomly assigned to the different conditions, comparisons were conducted for Experiment 1 (whole-face display) versus Experiments 4A (eye region alone) or 4B (mouth region alone), or 5A (face without eye region) or 5B (face without mouth region). Comparisons across expressions for a given facial region, and comparisons between each region and the whole-face condition, will reveal how important (either sufficient or necessary) a particular region is for an efficient detection of each expression. In Experiments 4A and 4B, we addressed the sufficiency criterion: If a region is sufficient for a visual search advantage, such region alone will produce the same effect as the whole face.

Method

Participants. Forty-eight psychology undergraduates (from 19 to 24 years of age) were randomly assigned to Experiments 4A or 4B (each with 24 participants; 19 women, 5 men).

Stimuli, design, procedure, and measures. The photographs of faces used in Experiment 1 were modified to be presented in Experiments 4A and 4B. Only the eye region (0.8° in height—21% of the whole face—× 3.0° in width) or only the mouth region (same size as the eye area) was used for Experiments 4A and 4B, respectively. Figures 6A and 6B illustrate how stimuli appeared in the different display conditions. The visual search arrays contained the selected region (either eyes or mouth) of six faces surrounding a central stimulus, all corresponding to the same model. The design, procedure, and measures were identical to those in Experiment 1 in all other respects.

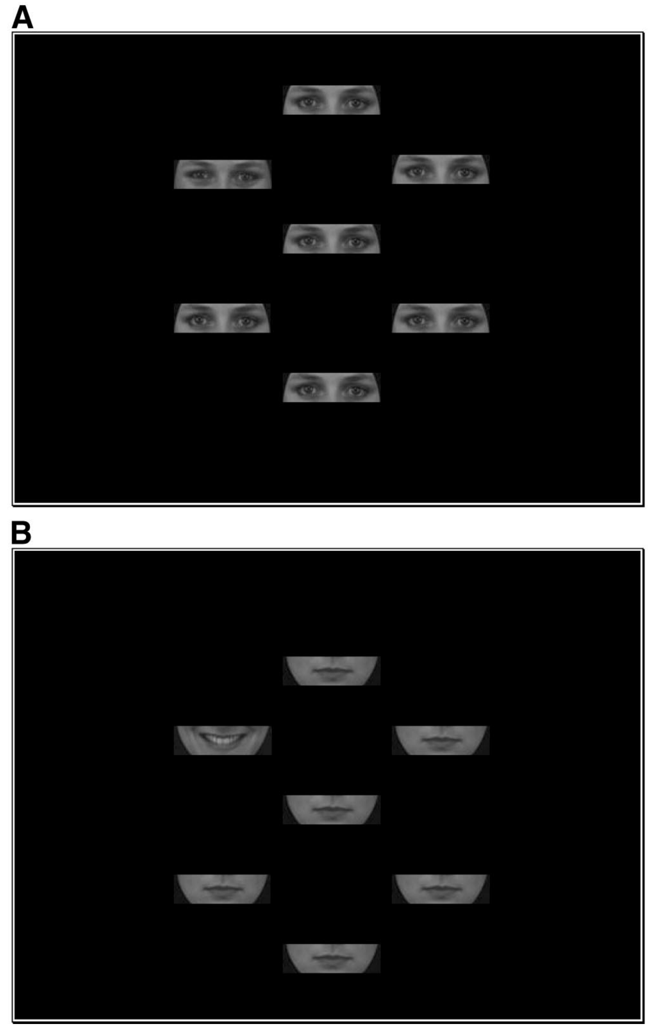

Figure 6. A (upper) and B (lower): Illustration of arrays of eye-only regions and mouth-only regions in Experiments 4A and 4B.

Results

Trials with all faces identical. When all faces in the display conveyed a neutral expression, a one-way (display: whole face vs. eye region vs. mouth region, i.e., Experiment 1 vs. 4A vs. 4B) ANOVA yielded no significant differences in response accuracy (.959 vs. .974 vs. .972, respectively; $F < 1$). Response times were affected by type of display, $F(2, 71) = 13.39$, $p < .0001$. Responses were faster for displays of whole faces ($M = 1{,}206$ ms) and of the mouth region ($M = 1{,}413$) than for displays of the eye region ($M = 1{,}866$).

Trials with one discrepant emotional target. Initially, the response accuracy and reaction time data of Experiments 4A and 4B were analyzed separately by means of 6 (expression) × 3 (location) ANOVAs. This served to make comparisons across expressions for the mouth and eye regions. Subsequently, response times for each region (i.e., Experiments 4A or 4B) were compared with those for the whole-face display (i.e., Experiment 1) in a 6 (expression) × 2 (type of display) ANOVA, with particular interest in the possible interactions. This served to determine the extent to which a facial region was sufficient to produce effects comparable with those of the whole face, depending on the type of expression. Mean response accuracy scores and response times are shown in Table 4 and Figure 7.

Experiment 4A (eyes visible only). Both response accuracy, $F(5, 115) = 50.19$, $p < .0001$, $\eta_p^2 = .69$, and reaction times were affected by expression, $F(5, 115) = 23.14$, $p < .0001$, $\eta_p^2 = .50$. As indicated in Table 4, accuracy was highest for angry, disgusted, and fearful targets, followed by sad and surprised targets, and it was poorest for happy targets. Similarly, as indicated in Figure 7, responses were faster for angry and disgusted targets than for fearful, surprised, sad, and happy targets.

The comparisons between the eye-only and the whole-face condition for reaction times (see Figure 7) yielded significant effects of expression, $F(5, 230) = 28.18$, $p < .0001$, $\eta_p^2 = .38$; display, $F(1, 46) = 50.08$, $p < .0001$, $\eta_p^2 = .52$; and an Expression × Display interaction, $F(5, 230) = 36.40$, $p < .0001$, $\eta_p^2 =$

Table 4
Mean Probability of Correct Responses in the Visual Search Task, as a Function of Type of Emotional Expression of the Target Face, in Experiments 4A (Eyes Only) and 4B (Mouth Only) and Experiment 1 (Whole Face; for Comparison)

	Type of expression displayed					
Face display	Happy	Surprised	Disgusted	Fearful	Angry	Sad
Eyes only						
M	.655 [c]	.819 [b]	.950 [a]	.891 [a]	.941 [a]	.826 [b]
SD	.154	.112	.058	.127	.087	.109
Mouth only						
M	.991 [a]	.981 [a]	.965 [a]	.964 [a]	.793 [b]	.648 [c]
SD	.021	.039	.044	.052	.091	.148
Whole face						
M	.981 [a]	.977 [a]	.962 [ab]	.932 [b]	.885 [c]	.867 [c]
SD	.037	.027	.061	.072	.084	.108

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent.

.19. Although search performance was slower for the eye region than for the whole face for all expressions (all post hoc *t*s > 3.40, $p < .001$; mean eyes only: 1,305 ms; mean whole face: 885 ms), the difference was greater for some expressions than for others. To examine the interaction, we analyzed reaction time difference scores between the eye region and the whole face (i.e., eye-region reaction times – whole-face reaction times) as a function of expression, and they were found to be influenced by emotional expression, $F(5, 115) = 36.55$, $p < .0001$, $\eta_p^2 = .61$. Difference scores were higher for happy expressions ($M = 706$ ms) than for any other expression (surprised: 515; fearful: 465; disgusted: 327; sad: 296; and angry: 209 ms), and they were higher for surprised and fearful expressions than for the other expressions. This means that, although the eye region is generally of little use in emotional face detection, its contribution—relative to the whole face—is particularly low for happy, surprised, and fearful expressions.

Experiment 4B (mouth visible only). The ANOVA yielded effects of emotional expression on response accuracy, $F(5, 115) = 98.56$, $p < .0001$, $\eta_p^2 = .81$, and response times, $F(5, 115) = 85.68$, $p < .0001$, $\eta_p^2 = .79$. As indicated in Table 4, accuracy was higher for happy, surprised, disgusted, and fearful targets than for angry targets, and it was poorest for sad targets. Similarly, responses were fastest for happy and surprised targets, followed by disgusted and fearful targets, and then by angry targets, and they were slowest for sad targets.

The comparisons between the mouth-only and the whole-face condition for reaction times (see Figure 7) yielded significant effects of expression, $F(5, 230) = 134.91$, $p < .0001$, $\eta_p^2 = .75$,

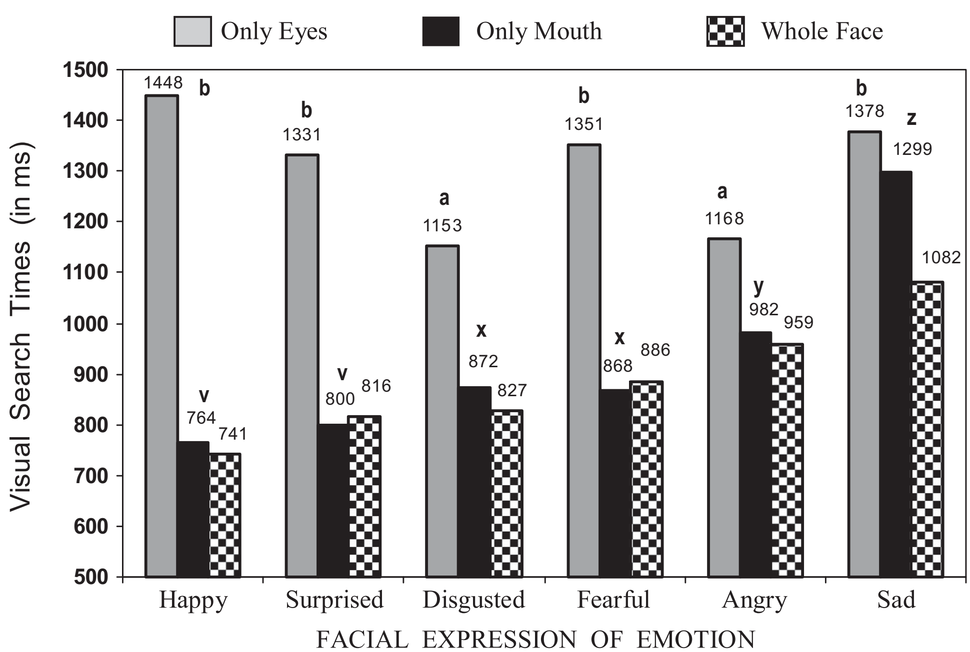

Figure 7. Mean response times for correct responses as a function of emotional expression in Experiments 4A (only-eye region) and 4B (only-mouth region) and Experiment 1 (whole face; for comparison). Significant differences in multiple contrasts are indicated by superscripts (a–b: only eyes; v–z: only mouth).

and an Expression × Display interaction, $F(5, 115) = 10.59$, $p < .0001$, $\eta_p^2 = .19$. The main effect of display was not significant ($F < 1$; mouth: 931 ms; whole face: 885 ms). Reaction times were longer in the mouth-only condition than in the whole-face condition for sad expressions, $t(46) = 2.89$, $p < .01$, whereas there were no significant differences between the two display conditions for the other expressions. This means that the mouth region is generally as effective as the whole face for detection, except for sad faces, in which the mouth makes a minimal contribution.

Discussion

In Experiments 4A and 4B, we examined the sufficiency criterion regarding the role of the eye and the mouth regions in emotional face detection. The mouth region alone was sufficient to yield an equivalent pattern of search differences between emotional expressions to that when the whole face was presented, with happy (and surprised) faces being searched most efficiently. In contrast, the eye region played a minor role in differentiating between facial expressions, and the happy face superiority disappeared when only this region was presented. Two prior studies in which the eye and the mouth regions of happy and angry faces were presented alone obtained equivocal findings. Fox and Damjanovic (2006) found that the angry eyes, but not the angry mouth, were detected faster than the corresponding happy eyes, in a context of neutral face regions. In contrast, Horstmann and Bauland (2006) found faster detection of the angry than the happy mouth, but no differences for the eye region, in a context of emotional (happy or angry) distractors. Apart from the fact that we found a happy rather than an angry face advantage (see our explanation in the General Discussion section), our findings are consistent with those of Fox and Damjanovic (2006) and Horstmann and Bauland (2006) in one important respect. In all three studies, single facial parts were sufficient to produce the same effect as the whole face. Thus, there is convergent support for a featural explanation of the superiority in emotional face detection using real face stimuli.

Experiments 5A and 5B

Role of the Eyes and the Mouth: Necessity Criterion

In Experiments 5A and 5B, we used a complementary approach to test the featural account of visual search performance by addressing the necessity criterion. If a face region is necessary for producing a detection advantage, the removal of such region from the whole face will eliminate the advantage of an emotional expression over others. This approach involved presenting the faces without either the eye or the mouth region, and comparing performance with the whole-face condition of Experiment 1.

Method

Participants. Forty-eight psychology students (from 19 to 24 years of age) were randomly assigned to Experiments 5A and 5B (24 participants each; 19 women, 5 men).

Stimuli, design, procedure, and measures. The face stimuli used in Experiment 1 were modified to be presented in Experiments 5A and 5B. Faces appeared without the eye (Experiment 5A) or the mouth (Experiment 5B) regions. The removed region was the same as that presented alone in Experiments 4A and 4B (subtending 3.0° × 0.8°). All other parts of the face were visible. Figures 8A and 8B illustrate how stimuli appeared in the two display conditions. The method was otherwise identical to that in Experiment 1.

Results

Trials with all faces identical. When all faces in the display conveyed a neutral expression, a one-way (face without eye region vs. without mouth region vs. whole face, i.e., Experiment 5A vs. 5B vs. Experiment 1) ANOVA yielded no significant differences in response accuracy ($F < 1$; $M = .971$ vs. .966 vs. .959, respectively). Response times were affected by type of display, $F(2, 71) = 6.02$, $p < .01$. Responses were faster for whole faces ($M = 1{,}206$ ms) than for displays without the mouth region ($M = 1{,}642$); response times for displays without the eye region ($M = 1{,}376$) were not significantly different from the others.

Trials with one discrepant emotional target. Initially, the response accuracy and latency data of Experiments 5A and 5B were

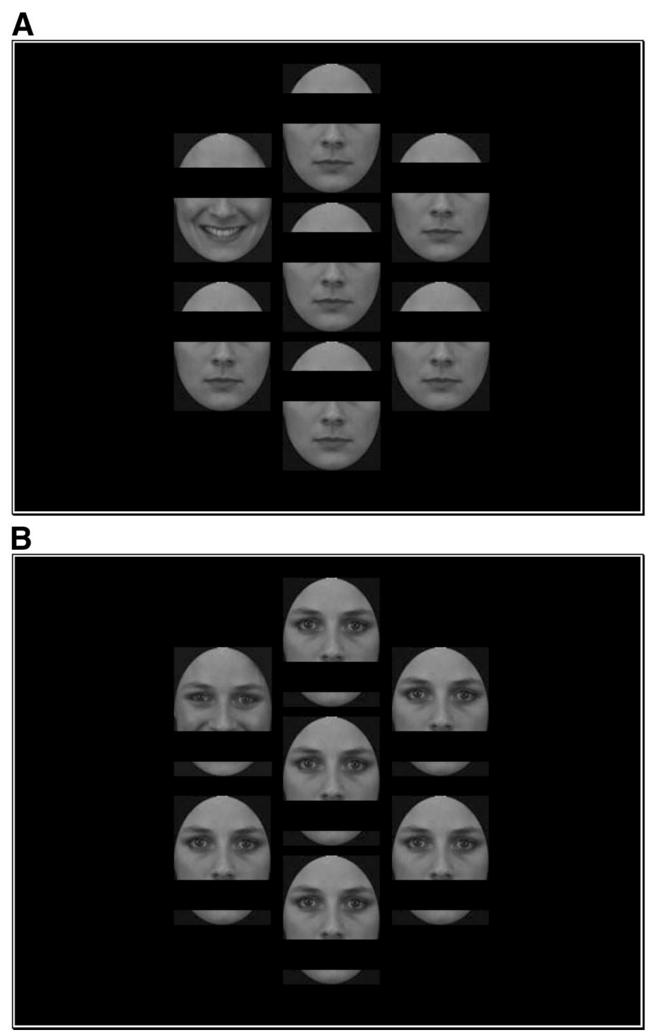

Figure 8. A (upper) and B (lower): Illustration of arrays of faces without the eye region and without the mouth region in Experiments 5A and 5B.

analyzed separately by means of 6 (expression) × 3 (location) ANOVAs. Subsequently, mean correct response times for each region (i.e., Experiments 5A or 5B) were compared with those for the whole-face display condition (i.e., Experiment 1) in a 6 (expression) × 2 (type of display) mixed ANOVA. Mean accuracy scores and reaction times are shown in Table 5 and Figure 9.

Experiment 5A (face without eyes). The ANOVA yielded significant effects of emotional expression on response accuracy, $F(5, 115) = 42.69$, $p < .0001$, $\eta_p^2 = .65$, and response times, $F(5, 115) = 63.15$, $p < .0001$, $\eta_p^2 = .73$. As indicated in Table 5, accuracy was higher for happy, surprised, disgusted, and fearful targets than for angry and sad targets. Responses were the fastest for happy targets, followed by surprised, disgusted, and fearful targets, followed by angry targets, and they were slowest for sad targets.

The comparison between the face-without-eyes and the whole-face condition for reaction times (see Figure 9) yielded significant effects of expression, $F(5, 230) = 114.06$, $p < .0001$, $\eta_p^2 = .71$. The effect of display ($F < 1$; mean face without eyes: 865 ms; mean whole face: 885) and the interaction, $F = 2.00$, $p = .10$, were not significant. This means that the absence of the eyes did not slow down the search of any emotional expression.

Experiment 5B (face without mouth). Expression affected response accuracy, $F(5, 115) = 10.19$, $p < .0001$, $\eta_p^2 = .31$, and latency, $F(5, 115) = 7.47$, $p < .001$, $\eta_p^2 = .25$. As indicated in Table 5, accuracy was highest for disgusted targets, followed by fearful, sad, angry, and happy targets, and it was poorest for surprised targets. Responses were fastest for disgusted targets followed by fearful, angry, and sad targets, and they slowest for happy and surprised targets.

The comparison between the face-without-mouth and the whole-face conditions for reaction times (see Figure 9) yielded significant effects of expression, $F(5, 230) = 20.64$, $p < .0001$, $\eta_p^2 = .31$; display, $F(1, 46) = 17.13$, $p < .0001$, $\eta_p^2 = .27$ (mean without mouth: 1,117 ms; mean whole face: 885 ms); and an Expression × Display interaction, $F(5, 230) = 20.90$, $p < .0001$, $\eta_p^2 = .31$. The interaction resulted from the fact that, for sad expressions, there was no significant difference (65 ms) in search time between the whole face and the face without mouth, whereas the difference was significant for all the other expressions. The extent to which performance decreased in the face-without-mouth condition relative to the whole-face condition varied as a function of expression, $F(5, 115) = 17.72$, $p < .0001$, $\eta_p^2 = .44$. Difference scores were greater for happy (426 ms) and surprised (378 ms) expressions than for fearful (196 ms), disgusted (172 ms), and angry (152 ms) expressions, which were greater than for sad expressions (65 ms). Thus the absence of the mouth was most detrimental for the detection of happy and surprised faces, and it was of little relevance for sad faces.

Discussion

The results of Experiments 5A and 5B indicate which face regions are necessary to produce a search advantage. The findings showed the important role of the mouth and the minor role of the eyes in accounting for search performance differences between facial expressions. The lack of the mouth region generally slowed down responses for all (except for sad) faces, relative to whole-face displays, whereas the lack of the eye region had a negligible impact. The detection of happy and surprised faces is especially dependent on the mouth region. Without the mouth region, not only did the search advantage of these faces disappear but detection times were generally longer than for the other expressions. In contrast, the eye region is not necessary for the detection of any emotional expression. The lack of the eye region did not significantly change the pattern of search differences in comparison with when the whole face was presented.

Experiment 6

Visual Saliency Attracts Attention to Smiles

From the previous experiments we can conclude that (a) there is consistent facilitated detection of happy faces, relative to other emotional faces, in a crowd of neutral faces, (b) this faster detection is due to the happy target faces selectively attracting first fixation and being localized earlier, and (c) this early attentional orienting is due to the higher visual saliency of happy faces. Complementary data indicate that (d) featural information from the faces, rather than their overall configuration, determines visual search performance, and (e) specific face regions, particularly the mouth, play a significant role. Putting all these findings together,

Table 5
Mean Probability of Correct Responses in the Visual Search Task, as a Function of Type of Emotional Expression of the Discrepant Face, in Experiments 5A (Without Eyes) and 5B (Without Mouth) and Experiment 1 (Whole Face; for Comparison)

	Type of expression displayed					
Face display	Happy	Surprised	Disgusted	Fearful	Angry	Sad
Without eyes						
M	.986 [a]	.972 [a]	.972 [a]	.964 [a]	.811 [b]	.814 [b]
SD	.032	.034	.036	.053	.097	.135
Without mouth						
M	.918 [b]	.866 [c]	.965 [a]	.934 [ab]	.922 [b]	.926 [b]
SD	.069	.078	.040	.068	.071	.061
Whole face						
M	.981 [a]	.977 [a]	.962 [ab]	.932 [b]	.885 [c]	.867 [c]
SD	.037	.027	.061	.072	.084	.108

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent.

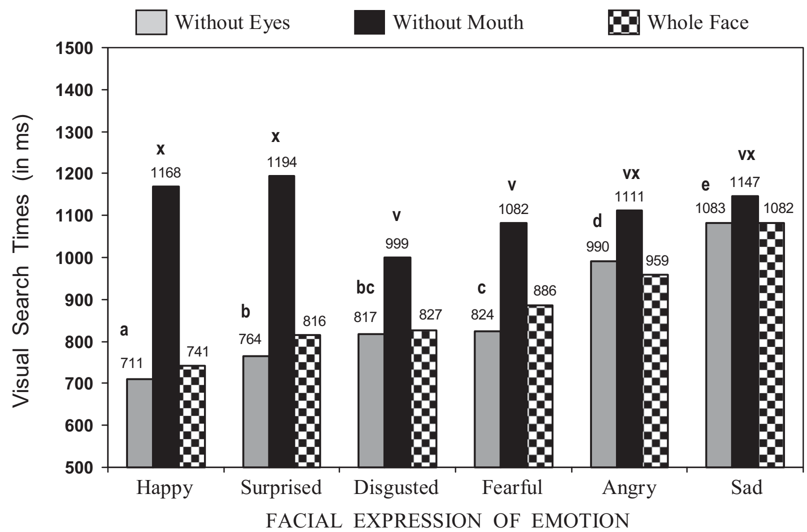

Figure 9. Mean response times for correct responses as a function of emotional expression in Experiments 5A (face without eye region) and 5B (face without mouth region) and Experiment 1 (whole face; for comparison). Significant differences in multiple contrasts are indicated by superscripts (a–e: without eyes; v–x: without mouth).

we can hypothesize that a visually salient, attention-capturing feature, such as the smile, can ultimately be responsible for the detection advantage of happy faces.

Using an integrative approach, in Experiment 6 we combined the saliency and the featural accounts of the detection advantage of happy facial expressions. We assessed the visual saliency of five regions (forehead, eyes, nose, mouth, and chin) of all expressions. One face was presented at a time parafoveally while the participants' eye movements were monitored during a facial expression identification task. Support for our hypothesis involves higher visual salience of, as well as more likely first fixation on, the mouth region of a happy face relative to other expressions and face regions.

Method

Participants. Twenty-four undergraduate psychology students (from 19 to 23 years of age; 25 right-handed; 19 women, 5 men) participated for course credit.

Stimuli. In addition to the 28 KDEF models used in Experiment 2, two more models were included (woman: no. 33; man: no. 22), each posing one neutral and six emotional expressions. The size of each face was increased to allow us to accurately determine the location of the fixations on different face regions. Each face subtended a visual angle of 8.4° (height) × 6.4° (width) at a 60-cm viewing distance, and appeared against a dark background. For the purposes of the data analysis, each face was segmented into five areas of interest (similarly to what we did in Experiments 4 and 5), although the whole face was visible during the experiment, and observers could not notice the separation between the areas. Vertically, the visual angles covered by each region were as follows: forehead (1.8°), eyes (1.6°), nose (1.8°), mouth (1.6°), and chin (1.6°).

Apparatus and procedure. The apparatus was the same as in Experiment 2. Each participant was presented with 210 trials in three blocks, randomly. A trial began with a central drift correction white circle (0.8°). A *prime period* started when the participant fixated the circle, which was followed by the presentation of a *single* face either to the left or right for 1 s. The distance from the center of the initial fixation circle and the inner edge of the face was 3°, so a saccade was necessary to bring the face to foveal vision. This was important to obtain eye movement measures. Following the 1-s face display, in a *probe period*, a word (*neutral*, *happy*, *angry*, *sad*, *disgusted*, *surprised*, or *fearful*) replaced the central circle—while the face remained visible—until the participant responded. The task involved pressing one of two keys to indicate whether the probe word represented the facial expression.

Design. There were two within-subjects factors: facial expression (neutral vs. happy vs. angry vs. sad vs. disgusted vs. surprised vs. fearful) and region (forehead vs. eyes vs. nose vs. mouth vs. chin) of the target face. On half of the trials, each participant was presented with a prime face that corresponded to the probe word (e.g., happy face and *happy* word); on the other half, the face and the word were different in content. Each participant was presented with the same facial expression of the same model only once, either on the left or the right visual field.

Measures. Visual saliency for each of the five predefined regions of each face was computed with the iNVT (Itti & Koch, 2000). The participants' eye movements were monitored to assess attentional orienting. This was operationalized as the location of the first fixation, that is, the probability that the first saccade following the onset of the face landed on each of the five regions. To further examine the initial distribution of fixations on each region, we also measured the number of fixations during the prime period. To determine whether the effect of saliency extended beyond the initial encounter with the face stimulus, we computed the number of fixations during the probe period. Recognition performance was assessed by the probability of correct responses and by reaction times from the onset of the probe word.

Results

Recognition accuracy and correct reaction times were analyzed by a 7 (expression) one-way ANOVA. A 7 (expression) $\times$ 5 (region) repeated-measures ANOVA was conducted on saliency scores, the probability of first fixation, and the number of fixations. Mean scores and significant multiple contrasts are shown in Table 6 and Figures 10 and 11.

Facial expression affected response *accuracy*, $F(6, 138) = 4.01$, $p < .01$, $\eta_p^2 = .15$, and *reaction times*, $F(6, 138) = 17.47$, $p < .0001$, $\eta_p^2 = .43$. Accuracy was higher for happy and disgusted faces than for fearful faces. Responses were fastest for happy faces, and they were slower for fearful faces than for all other emotional faces. In a lexical-decision task of a different experiment (24 participants), recognition times for the probe words alone were obtained, with no significant differences (happy: 630 ms; surprised: 633; disgusted: 632; fearful: 604; angry: 618; sad: 602; $F < 1$). This implies that the reaction time advantage of happy faces in the current experiment—as assessed by the corresponding probe words—is not attributable to differences in the processing of the words.

Saliency was affected by expression, $F(6, 174) = 52.18$, $p < .0001$, $\eta_p^2 = .64$; region, $F(4, 116) = 6.59$, $p < .0001$, $\eta_p^2 = .19$; and their interaction, $F(24, 696) = 29.41$, $p < .0001$, $\eta_p^2 = .50$. Separate one-way (expression) ANOVAs were conducted for the two regions of interest, that is, eyes and mouth. The nose region was also analyzed as a control condition, for comparison with the eye and the mouth regions. See the mean relative saliency scores of each of these areas within the whole face, for each expression, in Figure 10. The results for the forehead and chin regions are not reported because of their not being discriminative between faces, not providing much expressive information, and receiving practically no fixations (> 1%). For the eye region, an effect of expression, $F(6, 174) = 4.72$, $p < .001$, $\eta_p^2 = .14$, indicated that it was less salient in happy faces than in any other (except neutral, $p = .077$) type of faces. No differences as a function of expression occurred for the nose region. In contrast, the mouth region, $F(6, 174) = 24.95$, $p < .001$, $\eta_p^2 = .46$, was most salient in happy expressions, and it was more salient in disgusted and surprised faces than in neutral and sad faces.

For *probability of first fixation*, main effects of expression, $F(6, 138) = 6.56$, $p < .0001$, $\eta_p^2 = .22$, and region, $F(4, 92) = 28.39$, $p < .0001$, $\eta_p^2 = .55$, were qualified by an interaction, $F(24, 552) = 4.47$, $p < .001$, $\eta_p^2 = .16$. Separate one-way (expression) ANOVAs yielded significant effects for the mouth, $F(6, 138) = 10.28$, $p < .0001$, $\eta_p^2 = .31$, but not for the eye or the nose regions (see mean scores in Figure 11). The probability that the mouth was fixated first was higher for happy faces than for all others, except disgusted faces. The only additional significant contrast involved a more likely first fixation on the mouth of disgusted relative to neutral faces.

The *number of fixations* on the face during the *prime period* was not affected by expression, but it was by region, $F(4, 92) = 26.39$, $p < .0001$, $\eta_p^2 = .53$, and an Expression $\times$ Region interaction, $F(24, 552) = 11.72$, $p < .0001$, $\eta_p^2 = .34$. There were effects of expression in the eye region, $F(6, 138) = 9.75$, $p < .0001$, $\eta_p^2 = .30$; nose region, $F(6, 138) = 5.68$, $p < .001$, $\eta_p^2 = .20$; and mouth region, $F(6, 138) = 18.96$, $p < .0001$, $\eta_p^2 = .45$. As indicated in Table 6, the eye region was fixated less often in happy faces than the other faces; the same tendency occurred for the nose region. In contrast, the mouth region was fixated more often in happy faces than the other (except disgusted, $p = .074$) faces. The *number of fixations* on the face during the *probe period* was affected by expression, $F(6, 138) = 7.41$, $p < .0001$, $\eta_p^2 = .24$; region, $F(4, 92) = 24.85$, $p < .0001$, $\eta_p^2 = .52$; and their interaction, $F(24, 552) = 4.22$, $p < .0001$, $\eta_p^2 = .16$. Effects of expression occurred in the eye, $F(6, 138) = 8.76$, $p < .0001$, $\eta_p^2 = .28$, but not the mouth or the nose regions. As indicated in Table 6, the eye region was fixated less frequently in happy faces than the other (except disgusted, $p = .080$) faces.

Discussion

The mouth region was most salient in the happy faces, and this region was also first-fixated more often in the happy faces, in

Table 6
Mean Probability of Correct Responses, Total Correct Response Times (in Milliseconds), and Number of First Fixations on the Face During the Prime and the Probe Periods, as a Function of Type of Emotional Expression and Facial Region, in Experiment 6

	Type of expression						
Variable	Happy	Surprised	Disgusted	Fearful	Angry	Sad	Neutral
Accuracy							
M	.969 [a]	.939 [ab]	.951 [a]	.889 [b]	.941 [ab]	.925 [ab]	.913 [ab]
SD	.070	.066	.049	.080	.042	.066	.093
Response times							
M	701 [a]	794 [b]	804 [b]	923 [c]	800 [b]	847 [b]	866 [bc]
SD	179	241	233	270	215	220	265
No. of fixations prime period							
Eye region	0.50 [c]	1.08 [a]	0.82 [b]	1.07 [a]	0.87 [ab]	0.91 [ab]	0.98 [ab]
Nose region	0.71 [c]	1.01 [abc]	0.95 [bc]	0.99 [bc]	0.94 [bc]	1.20 [a]	1.11 [ab]
Mouth region	1.23 [a]	0.59 [c]	0.89 [b]	0.65 [bc]	0.84 [bc]	0.66 [bc]	0.64 [bc]
No. of fixations probe period							
Eye region	0.63 [c]	1.03 [ab]	0.89 [abc]	1.32 [ab]	0.94 [ab]	1.01 [ab]	1.07 [ab]
Nose region	0.91	0.97	1.09	1.03	1.00	1.11	1.08
Mouth region	0.77	0.71	0.74	0.88	0.72	0.68	0.74

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent.

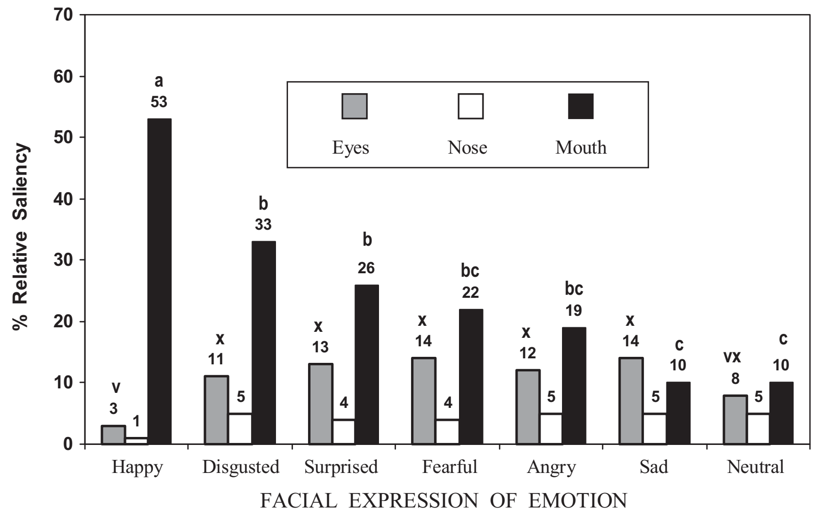

Figure 10. Mean relative saliency of each face region, as a function of emotional expression, in Experiment 6. Significant differences in multiple contrasts are indicated by superscripts (v–x: eyes; a–c: mouth).

comparison with most of the other faces.[1] This suggests that visual saliency of particular face areas determined the early selective direction of fixations to those areas within the face. The new findings of Experiment 6 regarding the comparisons between face regions add to those of Experiment 2 regarding the comparisons between different whole-face expressions. This role of saliency in attentional orienting is consistent with theoretical models (Itti, 2006; Itti & Koch, 2000). Our results regarding the number of fixations during the prime versus the probe period further confirm that saliency affects initial attention but not late processing. Differences between facial expressions in number of fixations during the prime period, that is, when the face was initially encountered, were related to visual saliency: More, or less, initial fixations occurred for the region that was more (i.e., the mouth of happy faces), or less (i.e., the eyes of happy faces), salient; in contrast, differences in number of fixations during the probe period were not related to saliency.

These findings are relevant to explaining the visual search advantage of happy faces that we have found in the previous experiments. We first noticed the fast localization of these faces and how this was related to their higher visual salience (Experiments 1 and 2). Next, support for a featural explanation was obtained (Experiment 3) and that the mouth region was the major source of detection differences (Experiments 4 and 5). We then hypothesized that some critical features at the mouth region could be particularly salient and attract attention early to this area, thus speeding up detection. The new findings of Experiment 6 support this hypothesis. Presumably, the critical, visually salient, attention-capturing feature is the smile. The smile has indeed been proposed as a key diagnostic cue in the recognition or identification of happy facial expressions (Adolphs, 2002; Leppänen & Hietanen, 2007), and it has been found that happy faces are identified faster than any other faces (Experiment 6; see also Calvo & Lundqvist, 2008; Palermo & Coltheart, 2004). This raises the question of whether a relatively high-level, meaningful feature such as the *smile*—which may be necessary for facial expression *identification*—is also involved in face *detection*. Alternatively, it is possible that lower level, perceptually based, and nonmeaningful components of the mouth shape are sufficient to account for orienting to and detection of happy faces. This more parsimonious view was held in the following experiment and reanalyses, in which we further explored the nature of saliency in the mouth region.

Experiment 7 and Reanalysis of Previous Data: The Nature of Saliency

Thus far our results have revealed that face detection is greatly dependent on visual saliency, and that the mouth region is the main source of difference between expressions. A major question is now concerned with *why* some faces are more salient than others, which results in earlier attentional orienting, which then facilitates localization and detection. The iNVT model (Itti & Koch, 2000) computes saliency from a combination of variations in three physical image properties: orientation, intensity, and color. The relative

[1] In some prior eye-movement studies that used singly-presented, nonemotional face stimuli (Henderson, Williams, & Falk, 2005) or prototypical expressions of all basic emotions (Adolphs et al., 2005), there were more fixations on the eye region than on any other region, including the mouth. Methodological differences can account for discrepancies between these data and ours. The face display time was 10 s (Henderson et al., 2005) and 5 s (Adolphs et al., 2005), instead of 1 s (current study). Eye movements can be affected by voluntary control in the long display conditions, but they are more subjected to automatic control by saliency in the short displays. Furthermore, the different effects of voluntary versus automatic eye movement control are likely to increase when the total *number of fixations* during the entire display period (Adolphs et al, 2005; Henderson et al., 2005) versus the probability of the *first fixation* (current study) is assessed. In any case, it should be noted that, in our study, the greater initial orienting to the mouth region occurred only for some emotional faces in which the mouth was especially salient. On average for all faces, and consistently with prior research, the number of fixations on the eye region was, in fact, greater (albeit nonsignificantly) than on the mouth region, both in the prime (M = 0.89 vs. 0.79, p = .36, *ns*) and the probe (M = 0.99 vs. 0.75, p = .23, *ns*) periods.

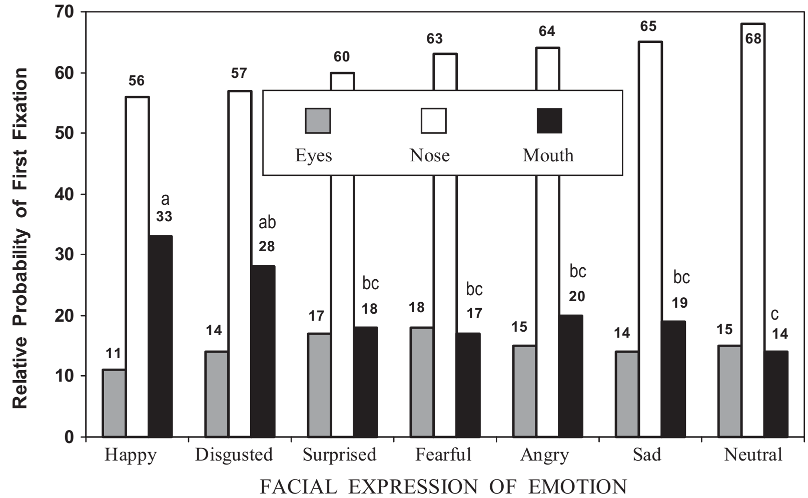

Figure 11. Mean relative probability of first fixation on each face region, as a function of emotional expression, in Experiment 6. Significant differences in multiple contrasts are indicated by superscripts (a–c: mouth).

contribution of each property is not specified by the model. It may be thought that intensity (or luminance) could be the major determinant of saliency. As applied to our own findings, this sounds plausible given that the main saliency source came from the mouth of happy faces, which typically involve a smile with visible *white* teeth. The exposed teeth could produce a blob of high local luminance, which would result in high local contrast around the mouth area. This would increase saliency, bias attentional orienting, and then facilitate detection.

Can the saliency of happy faces be reduced to local contrast and luminance caused by exposed white teeth? To examine this possibility, we used three approaches. First, we conducted Experiment 7, in which the faces were presented on white—instead of black—background displays. Second, we assessed the local luminance, local contrast density, and teeth exposure of the mouth regions of all faces. Third, we compared the faces with exposed teeth and those not showing teeth, regarding orienting and detection times.

Experiment 7

On a white background, the contrast of some white local features in a face, such as the teeth, will diminish. Accordingly, if the detection advantage of some faces (e.g., happy) is due to the white of some of their regions (e.g., mouth), such advantage will be significantly reduced with a white background, in comparison with when the faces are presented against a black background. In Experiment 7, the method was identical to that of Experiment 1 except that the faces were presented on a white background. Twenty-four new psychology undergraduates (from 19 to 23 years of age; 17 women, 7 men) participated in Experiment 7.

To determine whether the background modified the effect of type of emotional expression, we analyzed response accuracy and detection times by means of 6 (target expression) $\times$ 2 (black vs. white background) ANOVAs, thus combining the data from Experiments 1 and 7. Mean scores and multiple comparisons are shown in Table 7. For *response accuracy*, there was an expression effect, $F(5, 230) = 24.75$, $p < .0001$, $\eta_p^2 = .35$, a borderline background effect, $F(1, 46) = 3.23$, $p = .079$, $\eta_p^2 = .066$, but no interaction ($F < 1$). Accuracy was highest for happy, surprised, and disgusted targets, followed by fearful targets, and it was poorest for angry and sad targets. For *response times*, a significant effect of expression, $F(5, 230) = 75.54$, $p < .0001$, $\eta_p^2 = .62$, emerged, with no background or interactive effects (Fs < 1). Responses were fastest for happy targets, followed by surprised, disgusted, and fearful targets, which were detected faster than angry targets, and were slowest for sad targets.

The same pattern of detection differences between faces appeared in the white and the black background displays. The slightly poorer detection in the white condition ($M = .912$, accuracy; 929 ms) versus the black condition ($M = .934$, accuracy; 885 ms) may be due to the interference caused by the intense brightness of the background. The important point is that this interference affected all expressions similarly, with no interaction. These new results suggest that the detection advantage (and visual saliency) of some faces is not simply due to their having blobs or patches of high luminance or whiteness. This is also consistent with the absence of low-level luminance or contrast differences between whole-face expressions, reported in Experiment 1. Furthermore, if some faces are salient because of some *local* features, such as teeth, the salience of these features within the face is unlikely to change because of changes in the background (e.g., from black to white). The reason is that the saliency of a region or a feature is *relative to* the other parts of the face *within* which it appears, rather than the background. Accordingly, we next assessed the local luminance and contrast specifically for the mouth region, as well as the teeth area.

Assessment of Luminance and Contrast of the Mouth Region and Teeth Exposure

For each face stimulus of all the emotional expressions, we first assessed the presence versus absence of exposed teeth, as well as the pixels covered by the teeth (by means of Adobe Photoshop 6.0). In one-way ANOVAs (6: expression), the percentage of faces

Table 7
Mean Probability of Correct Responses and Reaction Times in the Visual Search Task, as a Function of Type of Emotional Expression of the Target Face, in Experiment 7, and the Black Background (Experiment 1) and White Background (Experiment 7) Displays Combined

	Type of expression					
Variable	Happy	Surprised	Disgusted	Fearful	Angry	Sad
			Experiment 7			
Accuracy (probability)						
M	.972 [a]	.948 [ab]	.955 [ab]	.911 [bc]	.854 [c]	.833 [c]
SD	.049	.075	.046	.084	.087	.144
Response times (in milliseconds)						
M	796 [a]	859 [ab]	888 [b]	911 [bc]	983 [c]	1,135 [d]
SD	180	204	188	219	197	201
			Experiments 1 and 7 combined			
Accuracy						
M	.977 [a]	.963 [a]	.958 [a]	.922 [b]	.870 [c]	.850 [c]
Response times						
M	769 [a]	837 [b]	857 [bc]	889 [c]	971 [d]	1,108 [e]

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent.

showing teeth, $F(5, 162) = 22.70$, $p < .0001$, $\eta_p^2 = .41$, and the mean size of the area covered by teeth, $F(5, 162) = 29.43$, $p < .0001$, $\eta_p^2 = .48$, were greater (all ps $< .05$) for happy faces than for the other expressions, and they were the least for sad expressions (see mean scores and multiple comparisons in Table 8). At first sight, this suggests that the saliency and the detection advantage of happy faces might be due to their generally having more luminance and contrast in the mouth region because of their white teeth; consistently, the disadvantage of sad faces could be due to their lack of exposed teeth.

To further examine this issue, we directly computed the luminance and contrast density (with Matlab 7.0) of the mouth region (as defined in Experiments 4–6) of all faces. In one-way ANOVAs, luminance, $F(5, 115) = 11.61$, $p < .0001$, $\eta_p^2 = .34$, and contrast, $F(5, 115) = 23.84$, $p < .0001$, $\eta_p^2 = .51$, varied as a function of emotional expression. Multiple comparisons indicated that for both luminance and contrast, the surprised mouth regions were generally the most different from the neutral faces, with the happy mouth regions being generally equivalent to most of the other emotional faces (see mean scores in Table 8). Accordingly, if the visible teeth contribute to saliency and visual search, it is not merely because of their luminance or contrast. Otherwise, the most salient and the fastest to be detected mouth regions (i.e., of happy faces) should have also been the ones with the greatest luminance and contrast, which was not the case. Conversely, teeth are rarely exposed in surprised faces, yet the luminance and contrast of their mouth region was the highest, and these faces generally enjoyed a detection advantage over most of the other expressions categories.

The Role of Teeth

We have shown that the faces and the mouths with more visible teeth (i.e., happy expressions) are the most salient, are especially likely to attract attention, and are detected faster—and that the reverse applies to faces with no exposed teeth (i.e., sad expressions). This suggests that the visual search advantage of some faces can be ultimately due to their displaying more teeth. However, we have also shown that a greater teeth exposure is not associated with greater luminance and contrast. This suggests that teeth contribute to salience, orienting, and detection not merely

Table 8
Mean Percentage of Faces (N = 28) With Exposed Teeth in Each Expression Category, Mean Size Area (in Pixels) Covered by Teeth (N = 28), Mean Size Area (in Pixels) Only for Faces Showing Teeth in Each Category (Variable N; see Percentage), and Mean Luminance and RMS Contrast Difference Scores Between Neutral Face and Each Emotional Face Stimulus for the Mouth Region

	Type of expression					
Variable	Happy	Surprised	Disgusted	Fearful	Angry	Sad
Exposed teeth (%)	96 [a]	21 [cd]	68 [b]	61 [bc]	32 [c]	0 [d]
Teeth area (all faces)	2,573 [a]	433 [cd]	1,262 [b]	1,334 [b]	709 [bc]	0 [d]
Teeth area (faces with teeth)	2,668 [a]	1,350 [b]	1,839 [ab]	2,049 [ab]	2,205 [a]	
Luminance	8.07 [bc]	14.17 [a]	9.81 [ab]	10.15 [ab]	7.91 [bc]	5.13 [c]
RMS contrast	0.025 [c]	0.079 [a]	0.028 [bc]	0.044 [b]	0.020 [c]	0.013 [c]

Note. Mean scores with a different superscript (horizontally) are significantly different; means sharing a superscript are equivalent. The number of pixels of the teeth area was obtained from a total face size of 99,824 pixels, which was identical for all faces within the oval-shaped window. RMS = root-mean-square.

because of the bright, white blob they produce. This raises the issue of why teeth can affect visual search, and whether teeth alone are sufficient produce the same effects in all facial expressions. It is possible that teeth yield such an advantage only when combined with specific surrounding facial features.

To address these issues, we grouped the stimulus faces according to whether they exposed teeth, separately for each expression category (see the percentage scores in Table 8). We then conducted F_2 (by-items) analysis by means of ANOVAs of expression by teeth exposure (yes vs. no) for visual search times in Experiments 1 and 2, and for the probability of first fixation on the target and localization times in Experiment 2. For *response times* in Experiment 1, there were effects of teeth, $F_2(1, 133) = 14.80$, $p < .0001$, $\eta_p^2 = .10$; expression, $F_2(5, 133) = 5.03$, $p < .0001$, $\eta_p^2 = .16$; and an interaction, $F_2(4, 133) = 5.01$, $p < .001$, $\eta_p^2 = .13$. Consistently, the respective effects on reaction times in Experiment 2 were as follows: $F_2(1, 133) = 36.14$, $p < .0001$, $\eta_p^2 = .21$; $F_2(5, 133) = 7.86$, $p < .0001$, $\eta_p^2 = .23$; and $F_2(4, 133) = 6.43$, $p < .0001$, $\eta_p^2 = .16$ (see the mean scores in Figure 12). Although detection time was generally shorter for faces showing teeth ($M =$ 810 and 785 ms, Experiments 1 and 2, respectively) than for those not showing teeth ($M = 996$ and 978 ms, Experiments 1 and 2, respectively), interestingly, this effect was qualified by the interaction. Separate contrasts between faces exposing versus not exposing teeth were conducted for each expression. Both for Experiments 1 and 2, the presence of teeth facilitated detection of happy and angry faces (all *p*s $< .0001$); in contrast, a similar tendency was nonsignificant for disgusted and fearful faces (all *p*s $\geq .10$), and the trend was opposite for surprised faces ($p < .05$, Experiment 1; $p = .40$, Experiment 2), in which teeth interfered with target detection; no comparison was possible for sad faces, as none showed teeth.

For the *probability of first fixation*, main effects of teeth, $F_2(1, 133) = 6.35$, $p < .025$, $\eta_p^2 = .046$; expression, $F_2(5, 133) = 6.22$, $p < .0001$, $\eta_p^2 = .19$; and an interaction, $F_2(4, 133) = 4.03$, $p < .01$, $\eta_p^2 = .11$, emerged. For *localization time*, there were main effects of teeth, $F_2(1, 133) = 18.96$, $p < .0001$, $\eta_p^2 = .13$; expression, $F_2(5, 133) = 5.33$, $p < .0001$, $\eta_p^2 = .17$; and an interaction, $F_2(4, 133) = 4.07$, $p < .01$, $\eta_p^2 = .11$. Faces with exposed teeth were more likely to be fixated first ($M = .489$ probability) and were localized earlier ($M = 380$ ms) than those without visible teeth ($M = .382$ and 480 ms, respectively). To decompose the interactions, we conducted separate contrasts between faces exposing versus not exposing teeth for each expression. Teeth increased the probability of first fixation and decreased localization time for happy, angry, and disgusted faces (all *p*s $<$ .05); in contrast, there was a similar but nonsignificant tendency for fearful faces ($p = .12$, first fixation; and $p = .17$, localization) and an opposite trend for surprised faces (*p*s $\leq .05$, first fixation; and $p = .29$, localization). The probability of first fixation scores are shown in Figure 12. The mean localization time scores (in milliseconds) for the teeth versus no-teeth faces, respectively, were as follows: happy (351 vs. 493), angry (389 vs. 558), disgusted (353 vs. 454), fearful (414 vs. 478), and surprised (391 vs. 365).

The interactions of expression and teeth exposure reveal that the influence of teeth on attentional orienting and detection efficiency is not uniform for all facial expressions. The facilitating effect of teeth was statistically significant only for some expressions; moreover, teeth tended to produce interference for others. This suggests that the role of teeth exposure varies as a function of the surrounding facial features, such as the shape of the mouth in which teeth appear. An alternative interpretation, however, is that the effect of teeth on visual search varies as a function of emotional expression simply because the size of the teeth area was greater for some expressions. To examine this alternative interpretation, in a one-way ANOVA we compared the size (in pixels) of the area covered by teeth for faces showing teeth in each expression category. This analysis is different from the one reported above (*Assessment of*

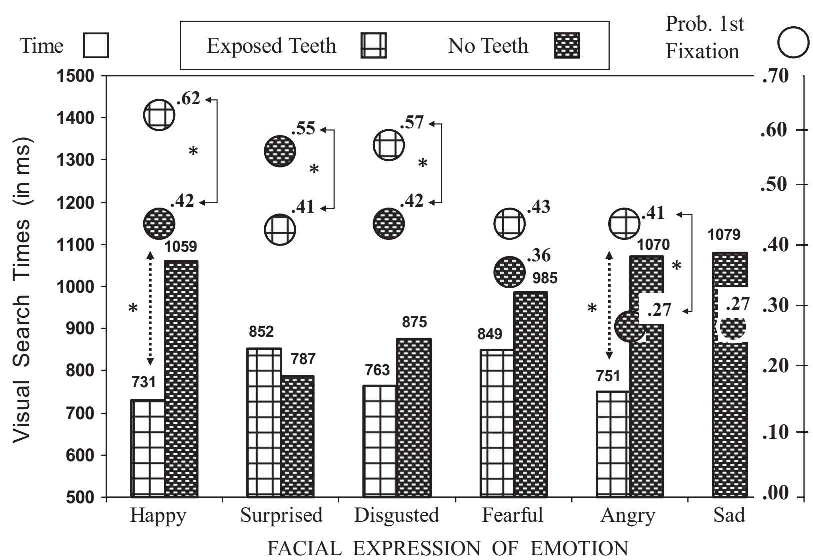

Figure 12. Mean response times and probability of first fixation on the target face, as a function of emotional expression and teeth exposure, averaged for Experiments 1 and 2 (response times). Vertical arrows and asterisks indicate significant differences between the faces with exposed teeth and the faces with no teeth visible.

Luminance and Contrast of the Mouth Region and Teeth Exposure) in that now only the means for the faces showing teeth are computed (rather than the means for all 28 faces of each category). An expression effect, $F(4, 77) = 6.34$, $p < .001$, $\eta_p^2 = .26$, followed by multiple post hoc comparisons, revealed that the size of the teeth area of happy faces was larger than that of surprised faces only, but the difference was not statistically significant with respect to the angry, fearful, and disgusted faces, with fearful and disgusted faces not differing significantly from surprised faces (see mean scores and contrasts in Table 8).

Conclusions

These results are generally in line with the hypothesis that the influence of teeth on visual search depends on the amount of teeth exposure. Thus, the greater teeth effects for the happy and the angry expressions could be due to their having larger teeth areas. However, the lack of statistically significant differences in most cases as well as the fact that teeth in the surprised faces can even interfere with performance (rather than simply being less facilitatory than for the other expressions) suggest that visual search differences between expressions are not merely due to quantitative differences in teeth exposure. This leaves room for the hypothesis that the effects of teeth partly depend on their combination with other specific surrounding facial features, such as the mouth *shape*. In fact, spatial *orientation* of the image components is one of the major factors underlying saliency, according to Itti and Koch's (2000) model. This is supported by the fact that there were significant visual search differences between faces that, otherwise, were not significantly different in teeth size. Probably, both hypotheses, that is, the amount of teeth alone and the combination of teeth with other features, are valid. Future research could try to disentangle their relative explanatory power.

At a more general level, the effects of saliency on attentional orienting and face detection that we have found are unlikely to be trivial. Saliency differences between facial expressions of emotion, and the corresponding and consistent orienting and detection differences, remained after having controlled for low-level confounds (i.e., luminance, contrast, energy, color, and texture). Rather, salience is sensitive to features that typically characterize emotional expressions, such as teeth, rather than merely artificial confounds. This should not lead us, however, to infer that salience reflects—and influences detection because of—the semantic or affective characteristics of the faces. Rather, saliency involves a combination of *physical* characteristics, with the salience of a single facial feature depending on other—probably local—features. The smile can thus facilitate attentional *orienting* and *detection* because of a large teeth exposure surrounded by upturned lip corners, rather than because it conveys a warm-hearted, friendly attitude to the viewer.

General Discussion

In this series of experiments, we investigated why some emotional facial expressions can be detected faster than others. The results of Experiment 1 revealed a visual search advantage for happy faces, followed by surprised, disgusted, and fearful faces, which were detected faster than angry faces, with performance being poorest for sad faces. In Experiment 2, the expressions that were detected faster were also more visually salient and more likely to be fixated first by human observers. Presumably, saliency attracted early initial orienting, which speeded up the detection process. In Experiment 3, the pattern of search differences remained even when the faces were presented upside-down. This suggests that the detection advantage is due to perception of prominent single features rather than to configural identification of expressions. In Experiments 4 and 5, this featural account was further explored by either presenting relevant face regions (mouth and eyes) alone or removing them from the face. The mouth made a strong contribution to visual search for most—especially, happy—expressions; the eyes played only a minor role for some expressions. Experiment 6 provided an integrative account of the saliency and the featural accounts. The happy mouth region was not only especially salient but also was most likely to receive the first fixation when faces were presented singly. This implies that happy faces are detected faster because the smile is a visually conspicuous feature that attracts attention reflexively. Finally, Experiment 7 and additional assessments of the mouth region indicated that saliency and its role cannot be reduced merely to blobs of white teeth but that it involves a combination of surrounding local features.

An Advantage in Face Visual Search as a Function of Emotional Expression

A consistent pattern of findings was replicated across various conditions in the current experiments: There was a visual search advantage for happy expressions, with faster detection—and, frequently, better response accuracy—than for others. This happy face superiority is in accordance with some previous findings that used photographs of real faces (Juth et al., 2005). It is, however, inconsistent with findings typically obtained with schematic faces (e.g., Calvo et al., 2006; Lundqvist & Öhman, 2005; Schubö et al., 2006) and with some studies that used real faces (Fox & Damjanovic, 2006; Hansen & Hansen, 1988; Horstmann & Bauland, 2006), in which an angry face superiority was found. The striking contrast between studies regarding the superiority of either happy or angry expressions needs to be explained.[2]

A possible explanation is that the faces used as experimental stimuli in some prior studies may not have been representative of

[2] It should be noted that in Hansen and Hansen's (1988), Fox and Damjanovic's (2006), and Horstmann and Bauland's (2006) studies, the face stimuli were drawn from the Pictures of Facial Affect (PFA) database (Ekman & Friesen, 1976), whereas, in Juth et al.'s (2005) experiments and the current study, KDEF stimuli (Lundqvist et al., 1998) were used. It might thus be thought that the empirical inconsistencies could simply be due to stimulus differences. This methodological account is, however, insufficient. First, Purcell et al. (1996) did not find an angry (or a happy) face advantage with PFA stimuli when the low-level confounds present in Hansen and Hansen's study were removed. Second, two studies adopting an individual differences approach also employed PFA pictures and reported results either only partially consistent (Gilboa-Schechtman et al., 1999) or nonconsistent (Byrne & Eysenck, 1995) with an angry-face superiority (see the introduction section). Gilboa-Schechtman et al. (1999) observed such superiority only for social-phobic participants. Byrne and Eysenck (1995) actually noted a happy face superiority for a low-anxious group. Third, Williams et al. (2005) used pictures from a different database (MacBrain Face Stimulus Set; Tottenham, Borscheid, Ellertsen, Marcus, & Nelson, 2002) and found an advantage in visual search of both angry and happy faces over other faces.

the natural variability of emotional expressions. Although facial happiness is consistently and universally characterized by a smile, there is considerable variability in the ways of expressing anger by different individuals and in different situations (e.g., visible teeth, lower lip depressed, lips tightly closed, frowning, outer brow raised, etc.; see Kohler et al., 2004). The high uniformity of the facial expression of happiness makes it very easy to recognize, whereas the angry expressions are more ambiguous and more often misjudged as neutral, disgusted, or even sad (Calvo & Lundqvist, 2008; Palermo & Coltheart, 2004). This implies that if a face stimulus sample is representative of the normal variability in real-life social interaction, happy faces will have an advantage because of a distinctive feature (i.e., smile) that can be used as an unequivocal diagnostic cue. In contrast, the greater featural variability of angry expressions would make them less easily discriminable. Thus, only if a small group of highly stereotypical exemplars with a prominent feature is used as stimuli, could the detection of angry faces equate or even be better than that of happy faces.

In accordance with this explanation, in all the studies supporting an angry real face advantage (Fox & Damjanovic, 2006; Hansen & Hansen, 1988; Horstmann & Bauland, 2006), the sample of stimuli was limited (sometimes, only two or three different models). In contrast, our sample (24 or 28 models) and that of Juth et al. (2005; 60 models) were considerably larger and thus more representative. The possibility that the angry expression advantage might be restricted to small selective subsets of facial stimuli was further corroborated by an item analysis of detection times that we conducted for the face stimuli used in our Experiments 1 and 2. Only for 5 models (women: no. 07 and no. 13; men: no. 17, no. 29, and no. 31) out of 28 was an angry face advantage found, whereas for the others there was generally a happy face advantage. Furthermore, for the five models showing an angry face advantage, the mean saliency of the angry face was greater than that of the happy face. This explanation can also be applied to schematic face stimuli, in that they are single prototypes of each expression, with unrealistic, exaggerated features. Particularly, schematic angry faces often have steep brows and/or a down-turned mouth, which probably attract attention because of their unusualness and enhanced saliency and, thus, facilitate search (see the Appendix).[3]

In addition to the detection superiority of happy faces, the current study makes a contribution regarding the comparisons of six different emotional expressions, for which a consistent pattern of findings also appeared. There was a superiority of surprised and disgusted faces over fearful and angry faces, which were all detected faster than sad faces. This extends the relevant comparisons beyond those allowed by previous studies, which included only two or three different expressions (Byrne & Eysenck, 1995; Fox & Damjanovic, 2006; Gilboa-Schechtman et al., 1999; Hansen & Hansen, 1988; Horstmann & Bauland, 2006; Juth et al., 2005; Purcell et al., 1996; Williams et al., 2005, included four emotional expressions). The fact that there were differences in visual search among most of the six facial expressions in our study represents an attractive theoretical challenge, which also calls for an explanation.

A Featural Account of Differences in Visual Search of Emotional Faces

According to a featural explanation, first, detection of discrepant target faces in a crowd is determined mainly by single features or parts of the target that make it discriminable from the distractors, rather than by configural information of the whole face. Second, the dependence on featural processing is greater for some expressions, particularly those in which prominent features appear consistently across most of the exemplars. On the contrary, detection of expressions without prominent features relies more on configural processing. Third, certain facial regions provide the most relevant features as reliable cues for search guidance and detection. The contribution of different facial regions will thus vary as a function of facial expression. This explanation was supported by data obtained with our spatially inverted arrays and the selective presentation of face regions.

The pattern of differences in search performance as a function of emotional expression was equivalent for upright and inverted displays. This allows us to infer that detection of facial expressions—in a normal upright orientation—relies mainly on local or featural information extracted from the faces. In two recent studies, researchers have addressed this issue by using photographic real faces (Fox & Damjanovic, 2006; Horstmann & Bauland, 2006), but they have provided divergent results. The advantage of angry over happy faces for upright displays disappeared (Fox & Damjanovic, 2006) or remained (Horstmann & Bauland, 2006) when the faces were inverted. Accordingly, different interpretations were offered: The emotional expression conveyed by the face (Fox & Damjanovic, 2006) or some visual feature (Horstmann & Bauland, 2006) was argued to be responsible for the detection advantage. Our own findings are consistent with those of Horstmann and Bauland (2006) in that the faces that showed an upright advantage (i.e., happy, in our case) maintained this advantage in the inverted condition as well. This suggests that the search superiority for any given expression is due to featural rather than to configural processing. Our results are also consistent with those of Fox and Damjanovic (2006) in that the detection of angry faces was in fact impaired by inversion. It is, however, important to note that, in our study, inversion slowed down reaction times for angry, fearful, and sad faces but not for happy, surprised, and disgusted faces. This reveals the relative involvement of featural (greater for happy, surprised, and disgusted faces) versus configural processing (considerable for sad, angry, and fearful faces).

[3] To further extend this explanation, we computed visual saliency for the schematic faces developed by Öhman et al. (2001). These face stimuli have been used frequently and have typically yielded an angry face detection advantage (Calvo et al., 2006; Horstmann, 2007; Juth et al., 2005; Lundqvist & Öhman, 2005; Mather & Knight, 2006; Tipples et al., 2002). Examples of these faces are shown in the Appendix, along with the basic saliency data. Essentially, an angry or a happy target face was presented among eight neutral faces in a 3 × 3 matrix, and saliency of the discrepant target was computed similarly to Experiment 2. Results revealed that saliency was greater for the angry than for the happy faces. In fact, the saliency values of the happy faces were equal to zero, thus indicating that they did not differ in saliency from the neutral context faces. The source for the greater visual saliency of the angry faces comes from the spatial incongruence between the contour of the face and the opposite orientation of the angry eyebrows and the angry mouth curvature (in contrast with the congruence in orientation for happy faces). These incongruences in spatial orientation would make the schematic angry faces highly salient and, hence, facilitate their detection.

The specific facial features that are relevant for rapid detection are located in the mouth region. The mouth alone—but not the eyes—was sufficient to produce the same pattern of differences in visual search as the whole face did. Consistent with this, removal of the mouth region nearly eliminated detection differences between expressions, whereas removal of the eye region did not. Research that used schematic faces has shown that no single feature (eyes, brows, and mouth) is sufficient to produce a search advantage, which occurs only when features are presented within whole-face configurations (Fox et al., 2000; Schubö et al., 2006; Tipples et al., 2002). Schematic features alone are, however, much less informative than regions of real faces. Prior research that used real face stimuli has obtained discrepant results, with one study supporting the role of the mouth (Horstmann & Bauland, 2006) and another favoring the importance of the eye region (Fox & Damjanovic, 2006), when comparing angry and happy faces. A wider range of expressions has allowed us to show that the importance of regions varies across expressions: The mouth is essential for the detection of happy and surprised faces; the eye region has some importance for disgusted and angry (and fearful) faces; for sad faces, neither region serves as an effective cue.

An Emotional Versus a Visual Saliency Explanation

From the previous section, we can conclude that the detection advantage of some emotional expressions is due to their fast featural processing, whereas others do not have distinctive features, and their recognition must thus rely to some extent on a slower configural processing. An important issue is whether the mechanism involved in the featural processing is purely perceptual, controlled mainly by bottom-up processing of the physical image properties of the face stimulus, or whether it may involve some top-down processing of the meaning conveyed by emotional expressions.

Having demonstrated the importance of certain facial features or regions for rapid detection of facial expressions, we can consider whether they are merely perceived as physical cues or whether they guide search because of their conveying the affective properties of the expressions with which they are associated. Lundqvist and Öhman (2005) argued that the high correlation between visual search performance and affective valence ratings of schematic faces, depending on whether different shapes of single features (eyes, brows, and mouth) are selectively included or not included in the face, is consistent with the affective processing view (see also Reynolds et al., in press). This implies that some features could be used as diagnostic cues that allow the observer to infer and identify the emotional expression of the face, without processing the whole face. The features could serve as a shortcut, or quick route, to categorize the associated expression (see Leppänen & Hietanen, 2007). However, other data do not support this view. Batty, Cave, and Pauli (2005) subjected geometrical shapes to aversive or neutral conditioning by associating them with threat-related or neutral pictures, respectively. When these shapes were presented later in a visual search task, the search slopes were similar for both the neutral and the threat-related targets. Thus, the association with threat did not lead to more efficient search. Although the association process and outcome may not be the same for facial features (e.g., teeth in a mouth with upturned lip corners) across prolonged real-life exposure as for abstract shapes in a constrained laboratory experience, the results of Batty et al. argue in favor of the perceptual view: For detection of emotional expressions (or any other visual target), what matters is the physical distinctiveness of the target rather than its affective meaning.

Our saliency data also support a perceptual explanation, as the higher visual saliency of happy faces was related to superior search performance. The iNVT algorithm (Itti & Koch, 2000) that we used for saliency computation assesses physical image properties, such as color, intensity, and orientation. The saliency map is thus obtained in a purely stimulus-driven or bottom-up manner (although it is technically possible to introduce top-down control in saliency mapping models; see Navalpakkam & Itti, 2005). Accordingly, no semantic or affective processing is involved, and the saliency weights do not reflect any meaningful representation, such as recognition of the object identity of the target. Nevertheless, the effects of physical saliency on human orienting can, of course, be modulated by contextual factors, such as task expertise or prior knowledge (Itti, 2006; Navalpakkam & Itti, 2005) or task-orienting instructions (see Underwood et al., 2006). Bottom-up visual saliency is thus not the only factor that guides human observers' attention, and the master saliency map in the human visual cortex must combine the saliency weights from the top-down and bottom-up saliency maps into an integrated, topographic representation of the relative behavioral relevance across visual space (Treue, 2003). At present, however, and given the strictly bottom-up saliency maps employed in our study, it is more parsimonious to consider the current data in line with the perceptual, bottom-up account, rather than with a semantic conceptualization, as the former account is sufficient to explain the behavioral results.

Conclusions: Integration of the Saliency and the Featural Accounts

There are consistent effects of emotional facial expression in visual search, with happy faces showing a special superiority and sad faces being at the greatest disadvantage. The mechanism responsible for such differences involves two stimulus factors, that is, facial features and their visual saliency, and two cognitive functions, that is, selective orienting versus facilitated decision. Conspicuous facial features—particularly in the mouth region—make some expressions—especially happy—visually salient. This attracts attention to them selectively and faster than to other emotional faces and regions. Because of this rapid localization of salient features, total detection time is shortened. Search efficiency is thus mediated by the direct effects of saliency on the early selective orienting of attention to facial features. In contrast, once a target face is localized, decisions about whether the target is different from the distractors, or about its identity, would not be affected by saliency.

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews, 1,* 21–62.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005, January 6). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433,* 68–72.

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

Byrne, A., & Eysenck, M. W. (1995). Trait anxiety, anxious mood, and threat detection. *Cognition and Emotion, 9,* 549–562.

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 527–551.

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

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

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods, 40,* 109–115.

Calvo, M. G., Nummenmaa, L., & Avero, P. (in press). Visual search of emotional faces: Eye-movement assessment of component processes. *Experimental Psychology.*

Carey, S., & Diamond, R. (1977, January 21). From piecemeal to configurational representation of faces. *Science, 195,* 312–314.

Cave, K. R., & Batty, M. J. (2006). From searching for features to searching for threat: Drawing the boundary between preattentive and attentive vision. *Visual Cognition, 14,* 629–646.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception and Psychophysics, 64,* 1004–1013.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception and Psychophysics, 65,* 352–358.

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia, 45,* 15–31.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

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

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion, 6,* 534–539.

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

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130,* 681–700.

Frischen, A., Eastwood, J. D., & Smilek, D. (in press). Visual search for faces with emotional expressions. *Psychological Bulletin.*

Gilboa-Schechtman, E., Foa, E. B., & Amir, N. (1999). Attentional biases for facial expressions in social phobia: The face-in-the-crowd paradigm. *Cognition and Emotion, 13,* 305–318.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54,* 917–924.

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 210–228.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33,* 98–106.

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

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

Itti, L. (2006). Quantitative modeling of perceptual salience at human eye position. *Visual Cognition, 14,* 959–984.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 1489–1506.

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.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research, 46,* 1762–1776.

Kohler, C. G., Turner, T., Stolar, N. M., Bilker, W. B., Brensinger, C. M., Gur, R. E., & Gur, R. C. (2004). Differences in facial expressions of four universal emotions. *Psychiatry Research, 128,* 235–244.

Latecki, L. J., Rajagopal, V., & Gross, A. (2005). Image retrieval and reversible illumination normalisation. *SPIE/IS&T Internet Imaging, VI,* 5670.

Leder, H., & Bruce, V. (1998). Local and relational aspect of face distinctiveness. *Quarterly Journal of Experimental Psychology, 51A,* 449–473.

Leppänen, J., & Hietanen, J. K. (2007). Is there more in a happy face than just a big smile? *Visual Cognition, 15,* 468–490.

Lewis, M. B., & Edmonds, A. J. (2005). Searching for faces in scrambled scenes. *Visual Cognition, 12,* 1309–1336.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces—KDEF [CD-ROM; ISBN: 91–630–7164–9]. Stockholm, Sweden: Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.

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

Mather, M., & Knight, M. R. (2006). Angry faces get noticed quickly: Threat detection is not impaired among older adults. *Journal of Gerontology: Psychological Sciences, 61B,* 54–57.

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

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception: II. Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9,* 431–444.

Müller, H. J., & Krummenacher, J. (2006). Visual search and selective attention. *Visual Cognition, 14,* 389–410.

Navalpakkam, V., & Itti, L. (2005). Modelling the influence of task on attention. *Vision Research, 45,* 205–231.

Nothdurft, H. C. (2006). Salience and target selection in visual search. *Visual Cognition, 14,* 514–542.

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.

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

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

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

Parkhurst, D., Law, K., & Niebur, E. (2002). Modelling the role of salience in the allocation of overt visual attention. *Vision Research, 42,* 107–123.

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.

Reynolds, M. G., Eastwood, J. D., Partanen, M., Frischen, A., & Smilek,

D. (in press). Monitoring eye movements while searching for affective faces. *Visual Cognition.*

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide.* Pittsburgh, PA: Psychology Software Tools.

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

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

Tipples, J., Atkinson, A. P., & Young, A. W. (2002). The eyebrow frown: A salient social signal. *Emotion, 2,* 288–296.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. (2006). Contextual guidance of eye movements in real-world scenes: The role of global features in object search. *Psychological Review, 113,* 766–786.

Tottenham, N., Borscheid, A., Ellertsen, K., Marcus, D. J., & Nelson, C. A. (2002). Categorization of facial expressions in children and adults: Establishing a larger stimulus set [Abstract]. *Journal of Cognitive Neuroscience, 14*(Suppl.), S74.

Treue, S. (2003). Visual attention: The where, what, how and why of visual saliency. *Current Opinion in Neurobiology, 13,* 428–432.

Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology, 18,* 321–342.

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

Appendix

Schematic Angry and Happy Faces and Their Saliency Values

Samples of neutral, angry, and happy schematic faces developed by Öhman, Lundqvist, and Esteves (2001) that were used in a number of studies showing a consistent angry face detection advantage (Calvo, Avero, & Lundqvist, 2006; Horstmann, 2007; Juth, Lundqvist, Karlsson, & Öhman, 2005; Lundqvist & Öhman, 2005; Mather & Knight, 2006; Öhman et al., 2001; Tipples, Atkinson, & Young, 2002; see also similar schematic faces—with an equivalent manipulation of two critical features, such as the eyebrows and the mouth—in Fox et al., 2000; Fox, Russo, Bowles, & Dutton, 2001; Schubö, Gendolla, Meinecke, & Abele, 2006) are provided here. For both the angry and the happy expressions, we computed the visual saliency (with the iLab Neuromorphic Vision C++ Toolkit; see Itti & Koch, 2000) of the four variants, depending on the shape of the eyes (A vs. B) and the length of the eyebrows (1 vs. 2). All four variants have been used in different studies, with comparable effects on detection. The expressive faces were embedded within a 3 × 3 matrix consisting of the target expression plus eight neutral faces, and the saliency of the discrepant face was computed similarly as in Experiment 2. Given such a small sample of items, no proper statistical comparisons could be performed. In any case, the results are clear-cut in that across five consecutive inhibition of returns, all the angry faces were highly salient, whereas the happy faces were not salient at all. The saliency values for each face variant are shown below the corresponding stimulus (see Figure A1).

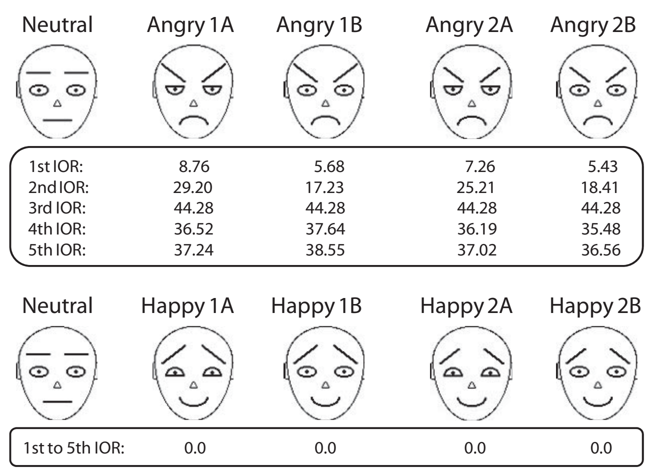

Figure A1. Schematic neutral, angry, and happy faces developed by Öhman, Lundqvist, and Esteves (2001) and mean visual saliency values of the angry and the happy targets in a crowd of eight neutral faces (3 × 3 matrices).

Received October 16, 2007
Revision received February 28, 2008
Accepted March 5, 2008 ■