

# Perceptual Grouping, Not Emotion, Accounts for Search Asymmetries With Schematic Faces

Stefanie I. Becker

The University of Queensland and University of Bielefeld

Gernot Horstmann

University of Bielefeld

Roger W. Remington  
The University of Queensland

Several different explanations have been proposed to account for the search asymmetry (SA) for angry schematic faces (i.e., the fact that an angry face target among friendly faces can be found faster than vice versa). The present study critically tested the perceptual grouping account, (a) that the SA is not due to emotional factors, but to perceptual differences that render angry faces more salient than friendly faces, and (b) that the SA is mainly attributable to differences in distractor grouping, with angry faces being more difficult to group than friendly faces. In visual search for angry and friendly faces, the number of distractors visible during each fixation was systematically manipulated using the gaze-contingent window technique. The results showed that the SA emerged only when multiple distractors were visible during a fixation, supporting the grouping account. To distinguish between emotional and perceptual factors in the SA, we altered the perceptual properties of the faces (dented-chin face) so that the friendly face became more salient. In line with the perceptual account, the SA was reversed for these faces, showing faster search for a friendly face target. These results indicate that the SA reflects feature-level perceptual grouping, not emotional valence.

*Keywords:* visual search, eye movement, threat capture, emotional faces, perceptual grouping

The number of stimuli that can be consciously perceived and responded to at any moment in time is severely limited. It is the role of attention to select potentially relevant stimuli for further processing. From the early beginnings of attention research, it has been clear that attention can be tuned to simple or basic properties of stimuli, such as size, color, and luminance (e.g., Wolfe, 1994). In visual search, target detection is very efficient when the target differs in a basic feature from the irrelevant distractors; detection times are unaffected by the number of irrelevant distractors, suggesting that attention is quickly allocated to the unique feature that signals the target. In cases where a search target cannot be distinguished from surrounding items on the basis of a basic feature, target detection is less efficient; detection times are slower and

increase systematically with the number of irrelevant background items, suggesting a serial search. As a general rule, search is usually inefficient and proceeds in a serial manner when the target is defined by a certain conceptual category membership, which is not correlated with a basic feature.

Potentially life-threatening stimuli—for example, snakes, spiders, and angry faces—constitute a category that appears to be an exception to this rule. According to the *threat capture hypothesis*, evolution has provided a second attentional guidance mechanism, one that constantly evaluates the environment for impeding danger to an individual's survival and well-being. When this threat detector is triggered, for instance, by a snake or a predator, ongoing behavior is immediately interrupted and attention is allocated to the threatening stimulus (as in the orienting response; e.g., Eastwood, Smilek & Merikle, 2001, 2003; Lipp & Waters, 2007; Lipp & Derakshan, 2005; Lundqvist & Öhman, 2005; Öhman, Lundqvist & Estevez, 2001; Rinck, Reinecke, Ellwart, Heuer, & Becker, 2005). According to the threat capture hypothesis, potentially threatening stimuli can hence capture attention automatically and independently of the goals and intentions of the observers.

A centerpiece of evidence gathered in favor of the threat capture hypothesis is the finding that a schematic depiction of an angry (threatening) face is found faster and more efficiently among schematic happy (friendly) faces than vice versa: a friendly face among angry distractor faces (e.g., Fox et al., 2000; Horstmann, 2007; Öhman et al., 2001). According to the threat capture hypothesis, the angry face can activate a threat detector preattentively, which can then guide attention to the threatening stimulus, even if attention is focused elsewhere in the scene. Hence, the

---

Stefanie I. Becker, School of Psychology, The University of Queensland, Brisbane, Australia and Department of Psychology, University of Bielefeld, Bielefeld, Germany; Gernot Horstmann, Department of Psychology, University of Bielefeld; Roger W. Remington, School of Psychology, The University of Queensland.

We thank David Cowan, Tess Hall, and Simanto Khandaker for assisting with the data collection, and Yuhong Jiang, Richard Piech, and Li Zhaoping for valuable comments on earlier drafts of the paper. This research was supported by a UQ Early Career Researcher Grant and an ARC grant (DP110100588) awarded to Becker, and an ARC grant (DP0666772) awarded to Remington.

Correspondence concerning this article should be addressed to Stefanie I. Becker, School of Psychology, McElwain Building, St Lucia QLD 4072, Australia. E-mail: s.becker@psy.uq.edu.au

search asymmetry (SA) for angry schematic faces is attributable to the fact that the threat detector facilitates detection and identification of a potentially threatening stimulus, (e.g., Eastwood et al., 2001, 2003; Hansen & Hansen, 1988; Öhman et al., 2001).

The threat capture account of the SA makes a number of predictions: (a) the threatening stimulus is found because of its emotional content, not because of a correlated perceptual feature; (b) the SA is determined by the threat index of the target (“target capture”), whereas the properties of the other, irrelevant items (“distractors”) are not causally relevant for the SA.

Both of these hypotheses, however, have been questioned. For instance, Fox, Russo, and Dutton (2002) proposed that threatening stimuli do not guide attention, but that it is more difficult to disengage or de-allocate attention from a threatening stimulus once it has been selected (see also Theeuwes, Atchley, & Kramer, 2000). According to this *de-allocation hypothesis*, angry faces do not exert their effect on a preattentive level and guide attention to a threatening target stimulus but on a postselectional level, after attention has been allocated to a threatening stimulus. Moreover, the de-allocation account claims that the SA is not driven by the identity of the target but by the identity of the distractors: search for an angry target among friendly distractors is faster than the reverse condition because attention dwells longer on the threatening distractor faces than on friendly distractor faces.

The view that the SA depends more on the identity of the distractors than on the identity of the target is supported by several observations: First, search for emotional faces usually produces set size effects consistent with an effortful, inefficient search, where distractors are selected and rejected individually. This would appear to argue against a target capture explanation of the SA, as inefficient search usually limits the potential impact of different targets (e.g., Nothdurft, 1993; Purcell, Stewart, & Skov, 1996). Second, search asymmetries are also observed on target absent trials, with search through a crowd of all-angry faces slowed compared to search among all-friendly distractors (e.g., Horstmann & Bauland, 2006; Horstmann, Scharlau, & Ansorge, 2006b). Third, when angry and friendly faces are presented among neutral distractors consisting of superimposing both types of faces, search was equally inefficient for angry and friendly face targets (Horstmann et al., 2006b). These findings, and especially the observation that search is also asymmetric in the absence of any targets, are difficult to explain by a target capture account such as the threat capture hypothesis, but easily accommodated by the de-allocation hypothesis.

## The Perceptual Grouping Account

The accounts discussed above both attribute the SA to the threat potential of angry faces and hence to emotional factors. In contrast to these *emotional factor explanations*, Horstmann, Heumann, and Borgstedt (2006a) proposed that the SA for angry faces may be attributable to perceptual factors. They noted that the faces used in visual search typically consist of a round or oval face outline, two dots representing the eyes, and an upward or downward curved mouthline, which determines the emotional expression of the faces (happy vs. angry; see Figure 1 for an example). In some studies, the faces additionally had a stroke or triangle representing the nose, and tilted strokes representing the eyebrows, with a

V-shaped configuration in angry faces (“\ /”) and an inverse V-shaped configuration in friendly faces (“/ \”); Horstmann, 2009).

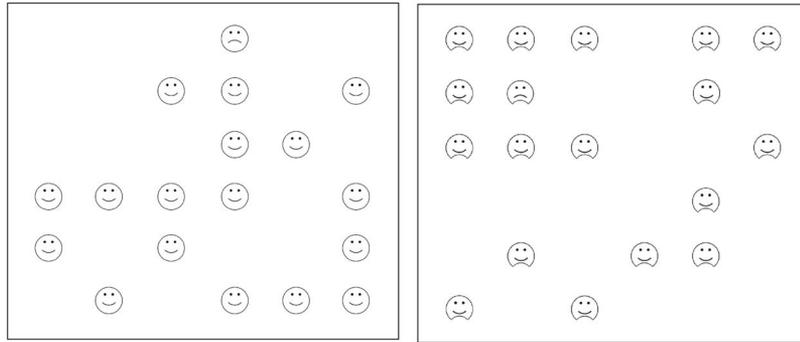
According to the *perceptual grouping account*, friendly faces are less perceptually salient than angry faces, which facilitates grouping of friendly faces when they constitute the distractors. In particular, the upward turned mouthline in friendly faces is more similar to the lower face outline (chin-portion), and this conformity or partial repetition of components diminishes local feature contrasts and reduces the number of distinct features in this region, so that friendly faces have a simpler Gestalt (see Horstmann et al., 2006a, 2006b; Horstmann, 2009). In comparison, the angry face is perceptually more salient, because the downward-turned mouthline creates a higher feature contrast with the lower face outline and adds a distinct feature to the face, which renders angry faces more complex. The same logic can also be applied to the tilted strokes representing the eyebrows: Because these are orthogonal to the face outline in angry faces, but tangential to the face outline in friendly faces, angry faces could also be more salient than friendly faces in virtue of the eyebrows. To note, we will use the term “saliency” to refer to these perceptual factors<sup>1</sup> only, which is not to be confused with other notions of saliency such as emotional or motivational saliency.

According to the perceptual grouping account (Horstmann et al., 2006a, 2006b), the perceptual properties of friendly faces facilitate *grouping*, so that a larger number of friendly faces can be selected and rejected in parallel when they constitute the distractors. Similar grouping accounts have been proposed to explain differences in search efficiency in other domains (e.g., search for letters, see Duncan & Humphreys, 1989). The core assumption of grouping accounts is that individual distractors are not selected in a strictly serial fashion, but in chunks, so that multiple items can be processed and rejected in parallel (as “structural units”; Duncan & Humphreys, 1989). Previous research has shown that the possibility and ease of grouping strongly depends on how similar the items are: For instance, grouping can encompass all items in the display when they all have the same color, whereas it is more difficult to group pink and red together, and altogether impossible to group red and green items (in virtue of their color). As an extension to this principle, the core idea of the perceptual grouping account is that, for intrinsically more complex stimuli such as faces, the perceptual organization of elements itself can render a stimulus more or less salient, which in turn can hamper or facilitate grouping.

The perceptual grouping account of the SA differs from the remaining views, in that (1) the SA is supposedly driven by perceptual and not emotional factors, (2) the SA is supposed to be mainly driven by distractor-based effects, which, however, operate at the preattentive level and not on the postselective level (e.g., of disengagement of attention). In line with the perceptual view,

<sup>1</sup> The core assumption of the perceptual grouping account is that grouping for friendly faces is facilitated because the search-relevant feature (i.e., direction of the mouthline) is more similar to the lower face outline in friendly faces than in angry faces. With this, friendly faces can be said to be less perceptually salient for a number of reasons, including that they (1) have reduced local feature contrasts, (2) lack a distinctive feature that is present in angry faces, and (3) have better Gestalt continuity (or a simpler Gestalt). All of these notions, and their implied mechanisms, are consistent with the perceptual grouping account.

A) Example of search displays (Experiments 1 and 2)



Examples of the masks used in Experiments 2A and 2B:



B) Schematic description of the conditions of Experiment 3.

	Faces				Scrambled Faces (Control)			
Standard								
Dented								
Instructed response, absolute mouth-dir.	down	up	down	up	down	up	down	up
Irrelevant mouth-dir. relative to face	down	up	up	down	—	—	—	—
	<i>compatible</i>		<i>incompatible</i>		<i>neutral</i>		<i>neutral</i>	

Figure 1. A, Examples of the stimuli in the target present displays of the visual search experiments, Experiments 1 and 2. B, The conditions of Experiment 3, where observers had to respond to the absolute mouthline-direction of a centrally presented single face. Processing of the irrelevant facial expression—or of the mouthline-direction relative to the face—will render the upright faces response-compatible and the inverted faces response-incompatible, allowing comparisons about the propensity to process the irrelevant emotional expression between dented and standard faces.

Horstmann, Becker, Bergmann, & Burghaus (2010) found that the SA for angry faces can be reversed by simply changing the face outline, leaving the emotional expression of the faces intact. In their study, the chin-portion of the schematic faces was curved inward, so that the mouthline of the angry face was parallel to the face outline, whereas the smiling mouthline of the friendly face was opposite to that of the chin. In line with the predictions of the perceptual account, search for a friendly dented-chin face among angry dented-chin faces was faster than vice versa, search for an angry dented-chin face among friendly dented-chin faces. A control experiment using standard schematic faces with the round (nondented) face outline moreover showed the typical SA for angry faces, demonstrating that the benefit for friendly dented-chin faces were not attributable to artifacts in the design.

To ensure that the dented-chin stimuli were still perceived as friendly and angry schematic faces, observers were also asked to rate the faces in terms of their emotionality. The ratings did not differ between dented and standard (i.e., nondented) faces, while angry faces were generally rated as being less pleasant than friendly faces. This indicates that the large differences in search for dented and standard faces were not attributable to differences in the perception of emotion (Horstmann et al., 2010).

The fact that the SA could be reversed by changing an irrelevant property of the face and without disturbing their emotional expres-

sion is inconsistent with the emotional account and presents strong evidence for the perceptual account. However, proponents of the emotional view could argue that subjective ratings cannot be taken as an indicator for the capability of stimuli to trigger the threat detector, as ratings presumably reflect higher order judgments and as such may not be reflective of the automatic and subconscious mechanisms that detect threat. Hence, it is still possible that the alterations to the chin region distorted the faces to such an extent that angry dented-chin faces failed to trigger the threat detector. In conclusion, the emotional view can be salvaged by claiming that the SAs for standard angry faces and friendly dented faces are mediated by two entirely different attentional systems: According to such a *two-systems view*, the SA for angry standard faces is determined by emotional factors, whereas dented faces fail to trigger the threat detector so that the SA for friendly dented faces is thus driven by entirely different, perceptual factors.

**Aim of the Study**

The aim of the present study was to test the different accounts (perceptual vs. emotional; grouping vs. de-allocation) by determining whether the SA is determined 1) by the identity of the background items or the target and 2) by perceptual factors (similarity) or by the emotional valence (threat index) of stimuli. To that aim,

we tracked the observers' eye movements during search for emotional faces, using both the standard stimuli and dented-chin stimuli (Experiment 1).

Eye movements allow a more in-depth analysis of the processes that finally determine RT and thus provide a richer source of information than RT (Zelinsky & Sheinberg, 1997; Zelinsky, Rao, Hayhoe, & Ballard, 1997). Monitoring the eye movements also allows separating search into different stages, allowing an analysis of the contributions from different factors in driving the SA (i.e., target capture, de-allocation, and distractor grouping; e.g., Becker, 2008a, 2008b, 2010; Becker & Horstmann, 2009).

Whereas Experiment 1 measured eye movements under conditions of free viewing, we used a gaze-contingent window technique in Experiment 2 to systematically manipulate the number of stimuli visible during a single fixation. This manipulation allows a more direct test of the grouping account, which predicts that the SA should be absent when only a single stimulus is visible (symmetrical search), whereas the SA should start to emerge and grow stronger with an increasing number of stimuli visible during a fixation.

In both Experiment 1 and 2, we used both the standard and the dented face stimuli to probe into the factors that drive the SA for angry standard faces and the SA for friendly dented faces. Note that the perceptual grouping account makes the rather strong claim that the SAs are mediated by the same perceptual factors and are based on the same underlying processes, whereas any emotional factor explanation has to assume that the SAs are based on entirely different mechanisms (threat index vs. perceptual factors). In the latter case, we can plausibly expect the eye movement behavior to differ in search for standard versus dented-chin faces; for instance, we could expect that the SA for standard faces is driven by target capture and/or difficulties to disengage attention, whereas the SA for dented faces is driven by grouping. By contrast, the perceptual grouping hypothesis would predict that both SAs are based on the same factors, so that the results for different eye movement parameters in Experiments 1 and 2 should be exactly mirror-reversed in search for standard and dented faces.

The prospective failure to find any differences in search for standard versus dented faces can however be regarded only as indirect evidence for a perceptual versus emotional factors explanation (first, because the conclusion is based on a null effect, and second, because it is always possible that different emotional vs. perceptual processes result in the same eye movement behavior). To test the emotional versus perceptual factors explanation more directly, Experiment 3 directly measured the extent in which standard and the dented stimuli are automatically processed as faces with a particular emotional expression.

### Experiment 1

In the first experiment, the SA for standard angry faces was compared to the SA of friendly dented faces while the eye movements of the observers were monitored. In different blocks, observers were asked to search either for a happy or angry face among all standard faces (smiley, frownie), or all dented faces. Dented faces differed from standard faces only by a dent in the chin region in the facial outline (see Figure 1 for an example of the stimuli). Observers were given no instructions with regard to their eye movements but had to manually respond to target presence

versus absence (50%). Of the available eye movement parameters, we report (1) the mean number of distractor fixations before selection of the target, (2) the mean dwell times, that is, the duration that the eyes remained fixated on a distractor, (3) the mean proportion of trials where the target was selected as the first item in the display, and (4) the latencies of these eye movements, that is, the duration the eyes remained fixated on the central fixation cross after the onset of the trial, to the point in time where the first eye movement to a target or distractor was initiated. The predictions were as follows:

1. According to the threat capture hypothesis, the SA is attributable to capture by angry faces when they constitute the target. Hence, we would expect that the angry target can be selected with a higher probability as the first item in the display than a friendly target (e.g., Becker, 2008a, 2008b).

2. The de-allocation hypothesis attributes the SA to difficulties of disengaging attention from angry face distractors. In this case, we would expect that the duration that the eyes remain fixated on a stimulus ("dwell time") is longer for fixations on angry faces than for fixations on friendly faces. This holds because eye movements to a location are usually preceded by attention shifts to that location, so that the eyes cannot be moved away from a location as long as attention is still engaged there (e.g., Deubel & Schneider, 1996). Correspondingly, the de-allocation hypothesis would predict longer dwell times for angry-face distractors than friendly face distractors.

3. The grouping account assumes that friendly faces can be grouped and rejected more efficiently than angry faces when they constitute the distractors. According to this view, multiple items can be selected and processed simultaneously during a single fixation, and the number of items that can be processed during a single fixation is higher with friendly face distractors than with angry face distractors. In other words, the visual span, or the attentional window, is larger in size when the distractors consist of friendly faces than when they consist of angry faces. With respect to the eye movement behavior, the grouping account would predict that fewer fixations are needed to search through friendly face distractors than through angry crowds. Thus, the mean number of distractor fixations before selection of the target should be higher in search for a friendly target than in search for an angry target (because the conditions require search through an angry crowd vs. a friendly crowd, respectively).

It is important to note that all accounts would allow for small variations in other dependent variables than those listed in the predictions. For instance, facilitated grouping of friendly face distractors could also lead to slightly shorter dwell times on friendly distractors, and/or produce small advantages in selecting the angry target as the first item on a trial. However, insofar as the accounts differ in the mechanism that is centrally responsible for the SA, they make different predictions about the effect that *predominantly* drives the SA. Hence, the different accounts can be distinguished by assessing the relative contributions of effects (1) in the proportion of first target fixations (threat capture), (2) dwell times (de-allocation), and (3) the number of distractor fixations (perceptual grouping) to the SA as typically observed in the mean RT.

With respect to search among dented versus standard faces, the perceptual grouping account would additionally predict no differences in the eye movement patterns in search among dented and

standard faces, which would be expected to be exactly mirror-reversed. By contrast, the two-systems view, which holds that the SAs for standard and dented-chin faces are mediated by completely different attentional systems (i.e., emotional system vs. perceptual system), would allow dissociations between the eye movement behavior in search for dented versus standard faces.

## Method

**Participants.** Twelve students from the University of Queensland, Australia, took part in the experiment as paid volunteers (\$10). Seven of them were female, five male, and they had a mean age of 22.5. All subjects had normal or corrected-to-normal vision and were naive as to the purpose of the experiment.

**Materials.** An Intel Duo 2 CPU 2.4-GHz computer with a 17" LCD monitor was used to generate and display the stimuli and to control the experiment. Stimuli were presented with a resolution of  $1280 \times 1024$  pixels and a refresh rate of 75 Hz. A video-based infrared eye-tracking system was used (Eyelink 1000, SR Research, Ontario, Canada) with a spatial resolution of  $0.1^\circ$  and a temporal resolution of 500 Hz. Participants were seated in a normally lit room, with their head fixated by the eye tracker's chin rest and forehead support, and viewed the screen from a distance of 65 cm. For registration of manual responses, a standard USB keyboard was used. Event scheduling and reaction time measurement were controlled by Presentation software (*Neurobehavioral Systems*).

**Stimuli.** The search display consisted of 18 schematic faces that were randomly drawn on 36 positions of a regular  $6 \times 6$  matrix (viewing area:  $24.5 \text{ cm} \times 32 \text{ cm}$ ). The faces consisted of black outlines and were presented against a white background. All faces had a diameter of 2 cm and were created such that the mouthline and the eyes were each 0.4 cm away from the facial outline (see Figure 1 for an example of the stimulus displays). Friendly and angry faces differed in the direction of the mouthline only (upward vs. downward curved, respectively). Standard faces had a round facial outline, whereas the facial outline of the dented stimuli was caved inward directly below the mouthline.

**Design.** The experiment consisted of the  $2 \times 2 \times 2$  within-subjects conditions "stimulus type" (standard vs. dented face), "emotional expression" (friendly vs. angry), and "target presence" (present vs. absent). Standard and dented faces were presented in different blocks, and emotional expression was also blocked, with the order of blocks being counterbalanced across participants. Target presence and the position of the target (when present) varied randomly within a block, so that the target was, on average, present on 50% of all trials. Participants completed 140 trials per block, resulting in 560 observations, and participants were allowed a short rest after each block.

**Procedure.** Each trial started with the presentation of a small black fixation cross ( $0.25 \text{ cm} \times 0.25 \text{ cm}$ ). Participants had to fixate on the center of the cross, and the search display was presented when the gaze was within 50 pixels (1.5 cm) of the center of the fixation cross, for at least 500 ms (within a time-window of 3,000 ms). Otherwise, participants were calibrated anew (nine-point calibration), and the next trial started again with the fixation control.

Upon presentation of the stimulus display, participants were required to search the display for the target and to press the right mouse button if the target was present and the left mouse button

when it was absent. The stimulus display remained on screen until response and was immediately followed by a feedback display. The feedback consisted in the black printed words "Correct!" or "Wrong!" (in Arial Black, 13 pt.), which were presented centrally and remained on screen for 500 ms. After an intertrial interval of 500 ms, in which a blank white screen was presented, the next trial started with the presentation of the fixation cross (and the fixation control).

Before the experiment, participants were calibrated with a randomized nine-point calibration and were given written instructions about their tasks. Participants were given no instructions concerning their eye movements but were instructed to respond to the target as fast as possible without making mistakes. The first 30 trials on each block were practice trials and were excluded from all analyses. On average, it took 45 minutes to complete the experiment.

## Results

**Data.** In this and the subsequent experiment, eye movements were parsed into fixations and saccades using the standard parser configuration of the Eyelink, which defines an eye movement as a saccade when its velocity and acceleration exceed  $30^\circ/\text{s}$  and  $8,000^\circ/\text{s}^2$ , respectively. Fixations were assigned to particular stimuli by superimposing a regular grid over the stimulus matrix, so that a fixation was counted as a fixation on a particular stimulus when the gaze was within the rectangular region belonging to that stimulus. Fixations on empty regions were rare and are not reported in the present study. Trials where manual responses were faster than 400 ms or slower than 5,000 ms were classified as outliers and excluded from all analyses. In Experiment 1, the exclusion criterion removed only a single trial (0.01% of all data).

**RT.** A  $2 \times 2 \times 2$  ANOVA with the variables of "stimulus type" (standard vs. dented face), "emotional expression" (friendly vs. angry), and "target presence" (present vs. absent) on mean RT revealed a significant main effect of target presence ( $F(1, 11) = 115.1$ ;  $p < .001$ ;  $\eta^2 = .91$ ), a significant interaction between stimulus type and emotional expression ( $F(1, 11) = 77.8$ ;  $p < .001$ ;  $\eta^2 = .88$ ), and a significant interaction between target presence and emotional expression ( $F(1, 11) = 5.8$ ;  $p = .035$ ;  $\eta^2 = .34$ ). The three-way interaction was also significant ( $F(1, 11) = 23.3$ ;  $p = .001$ ;  $\eta^2 = .68$ ), reflecting that emotional expression modulated RT differently on target present and absent trials in search for the dented face [ $F(1, 11) = 99.2$ ;  $p < .001$ ;  $\eta^2 = .90$ ], but not in search for the standard face [ $F(1, 11) = 2.6$ ;  $p = .13$ ].

Separate analyses showed the classical SA for the standard stimuli: Search was faster for an angry target among friendly distractors than vice versa, both in target present and target absent trials (all  $F_s > 12.2$ ;  $p_s < .005$ ;  $\eta^2_s > .53$ ). By contrast, searching among dented faces produced the reverse SA of faster search for a friendly target among angry distractors than vice versa (all  $F_s > 40.3$ ; all  $p_s < .001$ ;  $\eta^2_s > .78$ ; see Figure 2).

**Errors.** The same  $2 \times 2 \times 2$  ANOVA computed over the mean error scores yielded a significant main effect of target presence [ $F(1, 11) = 62.2$ ;  $p < .001$ ;  $\eta^2 = .85$ ], a significant interaction between emotional expression and stimulus type [ $F(1, 11) = 30.7$ ;  $p < .001$ ;  $\eta^2 = .74$ ], and a significant three-way interaction [ $F(1, 11) = 27.5$ ;  $p < .001$ ;  $\eta^2 = .71$ ], which was

## Results of Experiment 1

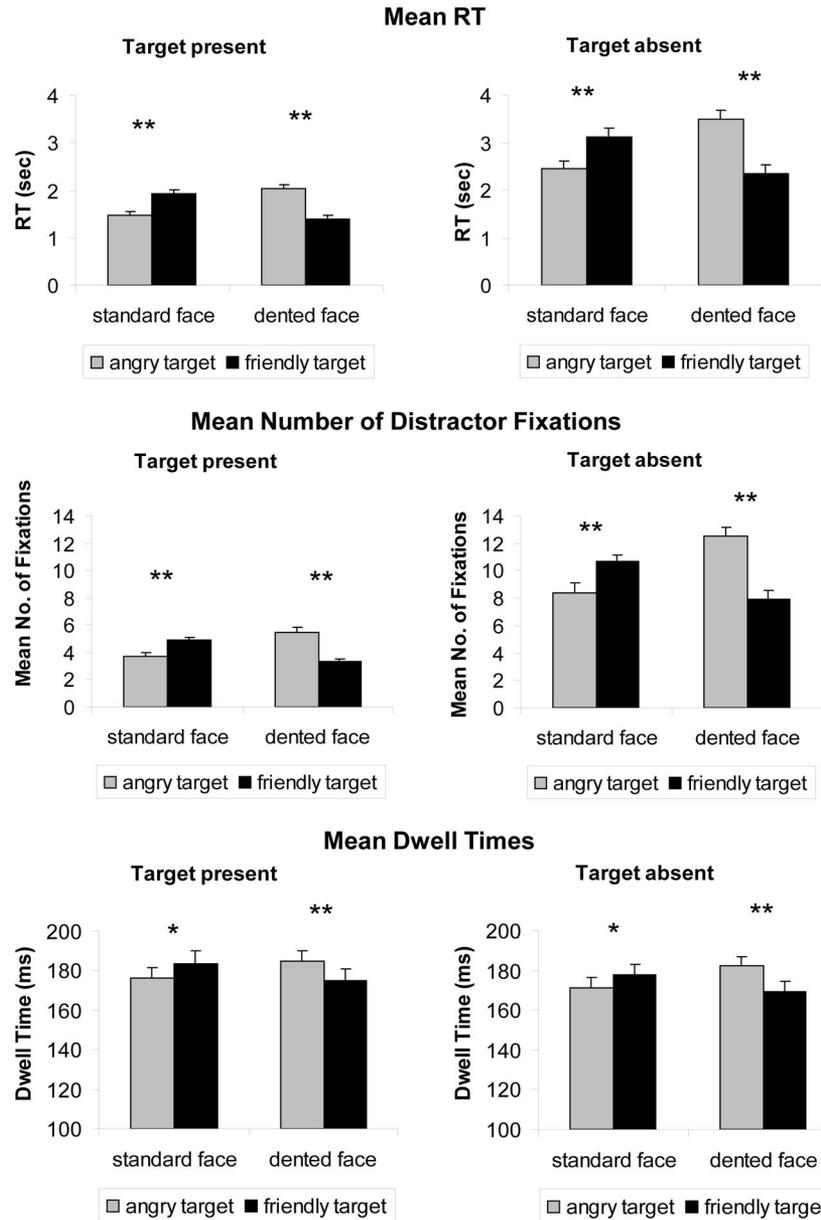


Figure 2. Mean RT, number of fixations before target selection, and mean dwell times for Experiment 1, where participants had to search for an angry or happy face that could have the standard or the dented face outline. In line with a perceptual account, the SA for angry standard faces was reversed for dented faces, with shorter RT, fewer distractor fixations and shorter dwell times in search for a friendly target. As discussed in the text, differences in the number of distractor fixations account for the major portion of the effects observed in the mean RT, consistent with a grouping account. Error bars represent the standard error of the mean (SEM). \*  $p \leq .05$ . \*\*  $p \leq .01$ .

attributable to the fact that the proportion of misses was disproportionately low in search for the dented friendly target (see Table 1A).

**Mean number of fixations.** The ANOVA computed over the mean number of distractor fixations showed that all main effects and interactions were significant, excepting the main effect of stimulus type. Analyzing only the mean number of fixations in

search for standard stimuli showed fewer fixations in search for an angry target among friendly distractor than vice versa [ $F(1, 11) = 18.3$ ;  $p = .001$ ;  $\eta^2 = .63$ ], which was highly significant on both target present trials [ $F(1, 11) = 33.4$ ;  $p < .001$ ;  $\eta^2 = .75$ ] and target absent trials [ $F(1, 11) = 11.9$ ;  $p = .005$ ;  $\eta^2 = .52$ ]. This result pattern was reversed for the dented stimuli, which showed

Table 1

## A. Mean Proportion of Manual Errors in Experiment 1

Exp. 1	Target present		Target absent	
	Standard face	Dented face	Standard face	Dented face
Manual errors				
Angry target	10.1% (2.0)	14.8% (2.3)	1.1% (0.7)	0.9% (0.6)
Friendly target	15.1% (2.3)	5.8% (1.1)	0.3% (0.2)	0.6% (0.2)
SA (friendly—angry)	5%	−9%**	−0.8%	−0.3%

## B. Proportion and Latencies of First Saccades to the Target Versus the Distractor in Experiment 1

	Standard face		Dented face	
	On target	On distractor	On target	On distractor
Angry target				
Proportion	8.9% (1.4)	80.7% (1.9)	5.7% (1.3)	86.6% (1.8)
Latency (ms)	305 (23)	241 (10.3)	291 (26.1)	240 (8.5)
Friendly target				
Proportion	4.3% (0.8)	87.0% (1.6)	11.4% (2.0)	80.4% (2.8)
Latency (ms)	279 (32.6)	256 (11.3)	250 (15.1)	231 (8.6)
SA (friendly—angry)				
Proportion	−4.6%*	7.7%*	5.7%*	−6.2%
Latency (ms)	−26	15	−41	−9

Note. (A) The mean manual errors and (B) the mean proportion of first saccades to the target versus the distractor in the target present condition of Experiment 1, depicted separately for search among standard versus dented faces, and conditions where the angry versus the friendly face was the target. The search asymmetry (SA) is given as the difference between search for friendly and angry faces, with \*  $p < .05$ ; and \*\*  $p < .01$ . Numbers in brackets denote the standard error of the mean (SEM).

significantly fewer distractor fixations in search for a friendly target among angry distractor faces than vice versa [ $F(1, 11) = 70.9$ ;  $p < .001$ ;  $\eta^2 = .86$ ; see Table 2]. Although the SA was significantly attenuated on target present trials (interaction term [ $F(1, 11) = 58.5$ ;  $p < .001$ ;  $\eta^2 = .84$ ]), it was highly significant on both target present and target absent trials [ $F(1, 11) = 57.8$ ;  $p < .001$ ;  $\eta^2 = .84$  and  $F(1, 11) = 72.5$ ;  $p < .001$ ;  $\eta^2 = .87$ , respectively].

**Dwell times.** The mean dwell times on the distractors—that is, the duration that the eyes remained fixated on a distractor—are depicted in Figure 2. Mimicking the results from the RTs and number of fixations, there was a significant interaction between emotional expression and stimulus type ( $F(1, 11) = 21.5$ ;  $p = .001$ ;  $\eta^2 = .66$ ), reflecting that, in search among standard faces, dwell times were shorter on the friendly distractors than on angry distractors, whereas this result pattern was reversed for the dented stimuli. The dwell time differences were rather small, ranging between 6 ms and 7 ms in search for standard faces, and between 10ms and 13ms in search for dented faces, but the differences were significant across all conditions [standard faces:  $F(1, 11) = 5.1$ ;  $p = .044$ ;  $\eta^2 = .32$  on present trials;  $F(1, 11) = 10.1$ ;  $p = .009$ ;  $\eta^2 = .47$  on absent trials; dented faces:  $F(1, 11) = 6.2$ ;  $p = .030$ ;  $\eta^2 = .36$  on present trials and  $F(1, 11) = 32.9$ ;  $p < .001$ ;  $\eta^2 = .75$  on absent trials; see Figure 2].

Dwell times were moreover longer on target present trials than on target absent trials [ $F(1, 11) = 19.8$ ;  $p = .001$ ;  $\eta^2 = .64$ ], which probably reflects a shortening of dwell times as overall search time increases.

**Mean proportion of first saccades to target and latencies.** The mean proportion of first saccades to the target versus the nontargets and the latencies of these saccades are depicted in Table

1B. To assess possible contributions of target-related effects (“capture”) to the SA, a  $2 \times 2 \times 2$  ANOVA was computed over the proportion of trials where the eyes selected the target as the first item. The analysis showed a significant stimulus type  $\times$  emotion interaction [ $F(1, 11) = 13.7$ ;  $p = .003$ ;  $\eta^2 = .56$ ], reflecting that, in search among standard faces, the angry face target could be selected as the first item in the display on a higher proportion of trials than the friendly face target [by 4.6%;  $F(1, 11) = 11.7$ ;  $p = .006$ ;  $\eta^2 = .52$ ]. By contrast, in search for dented-chin faces, the proportion of first eye movements to the friendly face target was significantly higher (by 5.7%) than to the angry face target [ $F(1, 11) = 6.1$ ;  $p = .031$ ;  $\eta^2 = .36$ ; see Table 1B].

The same analysis computed over the mean saccade latencies showed no significant effects between saccades to friendly versus angry target faces. Saccades to the target were initiated later than saccades to the distractors across all search conditions, but this difference was significant only for saccades to the angry standard face target [ $F(1, 11) = 10.1$ ;  $p = .009$ ;  $\eta^2 = .48$ ; see Table 1B].

## Discussion

The results from Experiment 1 replicated earlier findings that the SA for angry faces can be reversed with dented-chin faces, demonstrating that the reversal of the SA with dented-chin faces is a robust finding that can be obtained also in somewhat different conditions. More importantly, analysis of the eye movement data indicated that the SA was mainly attributable to the fact that more fixations were needed to find the friendly target. There were also significant differences in the proportion of first fixations on the target and the mean dwell time on angry faces, which were in the direction of the SA. However, these differences were too small to

Table 2  
*Experiment 2: Mean Number of Fixations and Dwell Times*

	Standard face			Dented face		
	Zoom 1	Zoom 2	Zoom 3	Zoom 1	Zoom 2	Zoom 3
<b>Target present</b>						
Number of fixations						
Angry target	8.3 (0.3)	6.6 (0.1)	4.9 (0.2)	7.8 (0.3)	7.6 (0.4)	6.4 (0.3)
Friendly target	7.9 (0.3)	6.9 (0.3)	5.9 (0.5)	8.2 (0.5)	6.9 (0.3)	5.4 (0.2)
SA ( <i>angry–friendly</i> )	0.4	−0.3	−1.0*	−0.4	0.7	1.1**
Dwell times						
Angry target	248 (0.6)	226 (5.8)	192 (3.0)	261 (9.4)	225 (6.3)	186 (7.8)
Friendly target	247 (6.4)	226 (6.2)	194 (3.9)	257 (9.8)	219 (6.4)	181 (5.8)
SA ( <i>angry–friendly</i> )	1	−1	−3	4	7	5
<b>Target absent</b>						
Number of fixations						
Angry target	16.5 (0.6)	13.5 (0.5)	10.4 (0.5)	16.3 (0.9)	15.1 (0.9)	13.1 (0.8)
Friendly target	16.8 (0.4)	13.8 (0.5)	11.3 (0.6)	16.3 (0.6)	13.9 (0.6)	11.3 (0.6)
SA ( <i>angry–friendly</i> )	−0.3	−0.3	−0.9	<0.1	1.3*	1.8*
Dwell times						
Angry target	248 (7.1)	224 (6.2)	188 (2.9)	254 (8.9)	224 (5.8)	187 (6.4)
Friendly target	247 (8.2)	227 (5.8)	190 (3.3)	253 (8.8)	219 (6.6)	177 (5.0)
SA ( <i>angry–friendly</i> )	1	−3	−2	1	6	11*

*Note.* The mean number of fixations and dwell times in Experiment 2, where the viewing area in search among standard faces (left) and dented faces (right) was restricted by gaze-contingent windows of different zoom factors. As the viewing area increased in size (zoom 1 to 3), the search asymmetry for angry standard faces and friendly dented faces began to emerge in the mean number of distractor fixations. Numbers in brackets denote the standard error of the mean.

\*  $p < .05$ . \*\*  $p < .01$ .

account for the effect in the mean RT. Immediate selection of the target was rare (<10% of all trials; see Table 1B) and occurred on only about 5% more of the trials when the target was angry than when it was friendly. This difference is too small to account for the large RT benefits in search for an angry target, of 442 ms and a 652 ms on target present and absent trials, respectively. This result would appear to argue against the threat capture view, in which the SA for standard angry faces is mostly attributable to immediate selection of potentially threatening angry faces. According to this view, we would have expected a higher proportion of first eye movements on both targets, more pronounced differences in the proportion of first fixations on angry versus friendly target faces, and an attenuation of the SA on target absent trials, contrary to the observed results.

The results are also at odds with the de-allocation hypothesis, as differences in dwell time amounted to only about 7 ms. With the average number of fixations being four fixations and 10 fixations on target present and absent trials, respectively, this difference is again too small to account for the SA in the RT, indicating that dwell-time differences cannot account for a major portion of the SA.

By contrast, the difference in the mean number of fixations was rather large, as search for a friendly target required 1.2 and 2.3 more fixations on target present and absent trials, respectively. These differences were large enough to account for the major portion of the SA observed in the mean RTs, as planning and executing an eye movement usually takes more than 200 ms (e.g., Findlay, 1997), and the average dwell time was about 180 ms. Even when we consider that these factors are not strictly additive in visual search (because the next eye movement is planned while the eyes are still fixating on a stimulus), it is clear that an additional fixation will produce costs in excess of 200 ms, so that the major portion of the RT differences in search for angry versus

friendly faces have to be attributed to differences in the number of distractor fixations.

The results thus support the grouping hypothesis that search through angry face distractors is more difficult because of less efficient grouping of angry faces, which reduces the number of stimuli that can be processed and rejected in parallel. The observed differences in the mean proportion of first eye movements could also be a side effect of differences in grouping: because, on average, a larger number of friendly face distractors can be grouped than of angry distractors, selecting the angry target with the first eye movement is slightly more probable than selecting a friendly target among angry distractors. Facilitated grouping could also lead to shorter dwell times on friendly distractors, because, in some instances, the spacing of the stimuli will limit the number of stimuli that are visible around the fixation, in which case facilitated grouping may be reflected in shorter dwell times. However, it is also possible that differences in the proportion of first fixations and dwell times are attributable to target and/or distractor effects that are independent of grouping (e.g., target capture or de-allocation).

More importantly, Experiment 1 provides some first evidence for a perceptual explanation of the SA: the results for the dented faces were exactly mirror-reversed compared to the results observed with the standard faces, which indicates that the SA for standard faces can indeed be reversed by changing some simple perceptual characteristics and without altering the emotional expressions of schematic faces. The SA for the dented friendly face was also driven mainly by differences in the mean number of fixations, and not by dwell time differences or differences in the ability to select the target as the first item on a trial (see Figure 2 and Table 1B). Hence, the results support

the view that both SAs were mainly attributable to more efficient grouping of the distractors.

## Experiments 2A and 2B

The results of Experiment 1 provide some first evidence for a perceptual account and specifically, for the view that the SA is attributable to differences in grouping. However, it could be argued that the results of Experiment 1 cannot yet provide compelling evidence for the grouping account, because the conclusion had to be based on comparing differences in eye movement data and RTs.

Experiments 2A and 2B were designed to provide a stronger test of the grouping account as an explanation for the SA for angry standard faces (Exp. 2A) and the SA for friendly dented faces (Experiment 2B). To that aim, we tested standard and dented stimuli using a gaze-contingent window paradigm, which revealed the identities of stimuli only within a region around fixation. In Experiments 2A and 2B, observers had to perform the same visual search task as in Experiment 1, but the identity of all search items was obscured by a mask, exempting only those items within a certain window that was continually centered on the observers' current gaze position. The size of the gaze-contingent window was varied across different blocks, so that observers could see either (1) only the fixated item (zoom 1 condition), (2) up to four search items simultaneously (zoom 2 condition), or (3) up to eight search items simultaneously (zoom 3 condition).

According to the grouping hypothesis, the SA should be absent in the narrow zoom condition, which did not allow grouping, and should only start to emerge when multiple items can be viewed simultaneously. The SA should also become stronger with further increasing the zoom factor, whereby search through friendly distractor crowds should profit more from a further widening of the window, resulting in the typical SA for angry target faces when the zoom factor is large.

By contrast, other accounts, such as the de-allocation hypothesis, would predict that the SA should be largely independent of the zoom condition, because attention should always linger longer on an angry face stimulus than on a friendly face. Hence, the SA should also be present in the narrow zoom condition, where only a single item is visible during the fixation.

## Method

**Participants.** Twenty-four students from the University of Queensland, Australia took part in the experiment as paid volunteers (\$10). Half of them searched for a standard face, and the other half searched for a dented face. All subjects had normal or corrected-to-normal vision and were naive as to the purpose of the experiment.

**Apparatus.** This was the same as in Experiment 2.

**Stimuli.** These were exactly the same as in the previous experiment, with the addition of masks that obscured the identity of stimuli: Masks were dials with diameter of 2.2 cm and resembled the steering wheel of a ship (see Figure 1, bottom).

**Design.** In separate blocks, observers searched either for a happy or angry face, with block order counterbalanced across participants. Within each block, target presence, target position,

and the zoom (viewing area around fixation) varied randomly. There were 3 zoom conditions; in zoom 1, 2, and 3, the visible region spanned 75 pixels (2 cm), 150 pixels (3.5 cm), and 300 pixels (7 cm) around the point they were fixating on, respectively.

## Results Experiment 2A: Standard Faces

**RT.** The mean RTs are depicted in Figure 2. A  $3 \times 2 \times 2$  ANOVA comprising the variables “zoom” (1 vs. 2 vs. 3), “emotional expression” (angry vs. friendly target) and “target presence” (present vs. absent) computed over the mean correct RT showed a significant main effect of the zoom factor [ $F(2, 22) = 155.2; p < .001; \eta^2 = .93$ ], target presence [ $F(1, 11) = 488.7; p < .001; \eta^2 = .98$ ], and a significant two-way interaction between both variables [ $F(2, 22) = 63.9; p < .001; \eta^2 = .85$ ]. The theoretically important two-way interaction between zoom factor and emotional expression was significant only on target present trials [ $F(2, 22) = 6.4; p < .007; \eta^2 = .37$ ] but not on target absent trials (see Figure 3).

Separate  $2 \times 2$  ANOVA probing for the SA within each zoom condition separately showed that emotional expression failed to affect mean RT in the zoom 1 condition ( $F < 1$ ) and modulated RT on target present trials in the zoom 2 condition but not on target absent trials (main effect of emotional expression: [ $F(1, 11) = 4.2; p < .066; \eta^2 = .85$ ; target present:  $F(1, 11) = 5.9; p = .033; \eta^2 = .35$ ; target absent:  $F(1, 11) = 1.8; p = .21$ ]. However, emotional expression strongly modulated RT in the zoom 3 condition [ $F(1, 11) = 9.1; p < .012; \eta^2 = .45$ ], both on target present and on target absent trials (both  $F_s > 6.4; p_s < .027$ ).

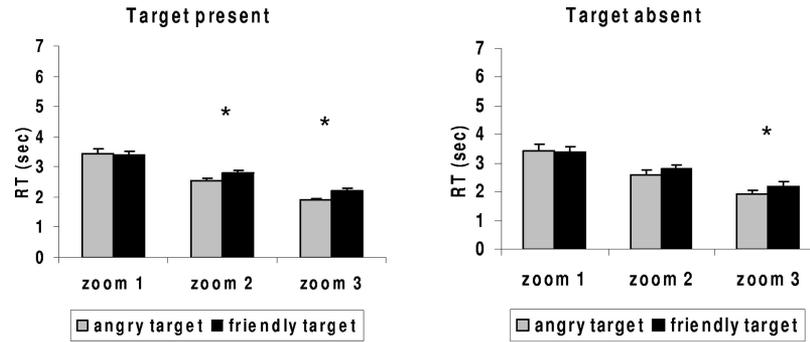
**Errors.** There were significant effects of emotional expression [ $F(1, 11) = 13.8; p = .003; \eta^2 = .56$ ] and target presence [ $F(1, 11) = 26.0; p < .001; \eta^2 = .70$ ] on the mean error scores. The two variables also interacted significantly with one another [ $F(2, 22) = 6.5; p = .027; \eta^2 = .37$ ], reflecting that observers frequently failed to detect the friendly target in the zoom 3 condition.

**Mean number of fixations.** The same omnibus ANOVA calculated over the mean number of distractor fixations showed similar results as found in the mean RT (see Table 2): There were main effects of the zoom factor [ $F(2, 22) = 108.4; p < .001; \eta^2 = .91$ ], target presence [ $F(1, 11) = 634.6; p < .001; \eta^2 = .98$ ], and a significant two-way interaction between both variables [ $F(2, 22) = 41.6; p < .001; \eta^2 = .79$ ]. The theoretically important interaction between zoom factor and emotional expression was also significant [ $F(2, 22) = 5.9; p = .012; \eta^2 = .35$ ], reflecting that the SA for angry face targets was absent in the lower zoom conditions and only started to emerge in the higher zoom conditions. In the zoom 1 and 2 conditions, emotional expression did not modulate the number of distractor fixations ( $F < 1$ ), whereas in the zoom 3 condition, significantly more distractor fixations were made in search for the friendly face than in search for the angry face [ $F(1, 11) = 5.7; p = .036; \eta^2 = .34$ ].

**Dwell times.** The mean dwell times on distractors—that is, the duration the eyes remained fixated on a distractor item—were also analyzed (see Table 2). The results showed that mean dwell times were affected by the zoom factor only [ $F(2, 22) = 112.4; p < .001; \eta^2 = .91$ ], reflecting that narrowing the zoom elongated dwell times.

## Results of Experiment 2

### 2A) Mean RT for Standard Faces



### 2B) Mean RT for Dented Faces

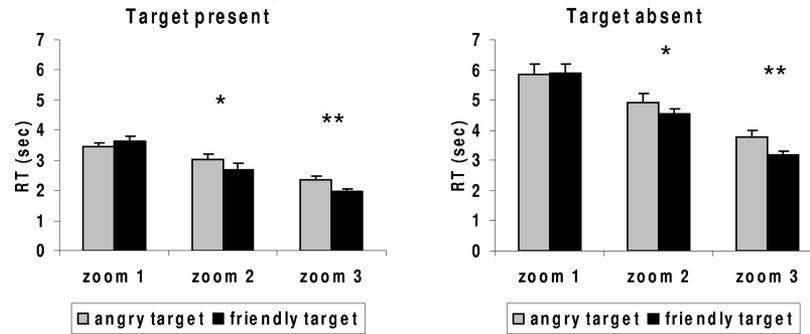


Figure 3. Mean RTs of the restricted-viewing Experiments 2A and 2B, depicted separately for search among standard faces (top panel) and dented faces (bottom panel), and separately for each of the different zoom conditions. As predicted by the grouping hypothesis, the search asymmetry for angry faces and the reversed search asymmetry for dented faces are absent when viewing is restricted to a single item and emerges and becomes stronger the more stimuli are visible during a single fixation. \*  $p \leq .05$ . \*\*  $p \leq .01$ .

### Results Experiment 2B: Dented Faces

**RT.** The omnibus ANOVA calculated over the mean RT showed a significant main effect of the zoom factor [ $F(2, 22) = 93.9$ ;  $p < .001$ ;  $\eta^2 = .89$ ], target presence [ $F(1, 11) = 179.9$ ;  $p < .001$ ;  $\eta^2 = .94$ ], a significant two-way interaction between both variables [ $F(2, 22) = 33.7$ ;  $p < .001$ ;  $\eta^2 = .75$ ], and, most importantly, a significant interaction between zoom factor and emotional expression [ $F(2, 22) = 12.0$ ;  $p = .002$ ;  $\eta^2 = .52$ ]. The interaction was due to the fact that the SA for friendly face targets was absent in the lower zoom conditions and only started to emerge in the higher zoom conditions. Separate ANOVAs computed over the RT in each zoom condition showed that RT did not differ between angry and friendly face targets in the zoom 1 condition ( $F < 1$ ), whereas in the zoom 2 and 3 conditions, RT for the friendly face target were significantly faster than for the angry face target [zoom 2:  $F(1, 11) = 5.2$ ;  $p = .044$ ;  $\eta^2 = .32$ ; zoom 3:  $F(1, 11) = 10.3$ ;  $p = .008$ ;  $\eta^2 = .48$ ].

**Errors.** The omnibus ANOVA calculated over the mean error scores showed a significant main effect of target presence [ $F(1, 11) = 13.0$ ;  $p = .004$ ;  $\eta^2 = .54$ ] and a significant interaction between zoom factor and emotional expression [ $F(2, 22) = 12.1$ ;  $p < .001$ ;  $\eta^2 = .52$ ]. The three-way interaction was also significant [ $F(2, 22) = 10.6$ ;  $p = .001$ ;  $\eta^2 = .49$ ], reflecting that the overall

trend to miss the friendly target more frequently than the angry target was reversed in the zoom 3 condition.

**Mean number of fixations.** The results from the mean number of fixations on the distractors are depicted in Table 2. As can be seen in the Table, the results exactly mirrored the results from the RT. The  $3 \times 2 \times 2$  ANOVA showed a significant main effect of the zoom factor [ $F(2, 22) = 92.9$ ;  $p < .001$ ;  $\eta^2 = .89$ ], target presence [ $F(1, 11) = 284.4$ ;  $p < .001$ ;  $\eta^2 = .96$ ], and a significant two-way interaction between both variables [ $F(2, 22) = 23.8$ ;  $p < .001$ ;  $\eta^2 = .69$ ]. The important interaction between zoom factor and emotional expression was also significant [ $F(2, 22) = 11.3$ ;  $p = .004$ ;  $\eta^2 = .51$ ], reflecting that the SA for friendly face targets was absent in the lower zoom conditions and only started to emerge in the higher zoom conditions. The mean number of distractor fixations did not differ between angry and friendly face targets in the zoom 1 condition ( $F < 1$ ), whereas in the zoom 2 and 3 conditions, there were fewer fixations in search for the friendly face target than for the angry face target [zoom 2:  $F(1, 11) = 5.9$ ;  $p = .032$ ;  $\eta^2 = .35$ ; zoom 3:  $F(1, 11) = 11.7$ ;  $p = .006$ ;  $\eta^2 = .52$ ].

**Dwell times.** The mean dwell times on the distractors showed only a significant main effect of zoom [ $F(2, 22) = 86.6$ ;  $p < .001$ ;  $\eta^2 = .89$ ], with dwell times becoming shorter when the zoom

factor increased and more stimuli became visible. The eyes also dwelt slightly longer on angry distractor faces than on friendly distractor faces, but this effect failed to reach significance [ $F(1, 11) = 3.7$ ;  $p = .079$ ;  $\eta^2 = .25$ ].

## Discussion

The results of Experiment 2A clearly showed that the SA for angry schematic faces occurs only when multiple distractors are visible during a fixation. Previous studies already showed that the SA for angry faces and SAs with other stimuli are typically absent when only a single item is presented in the display (set size 1 condition; e.g., Horstmann, 2007; Treisman & Gormican, 1988; Treisman & Souther, 1985). However, this result pattern is consistent both with a grouping account and a de-allocation account, because, with a single stimulus, there is no opportunity for grouping, but also no need to de-allocate attention to another stimulus after selection.

The present experiment tested search performance under conditions that required frequent de-allocation and reallocation of attention to other stimuli, while systematically varying the number of distractors that could be rejected in parallel. The results showed that, contrary to the de-allocation account, the SA was absent in the narrow zoom condition and increased when more stimuli were simultaneously visible. This indicates that the SA critically depends on the differences in the processing speed of multiple items, in line with the grouping account. Hence, Experiment 2A provides strong support for the view that the SA for angry schematic faces is caused by difficulties to group (more salient) angry faces when they constitute the distractors.

Experiment 2B moreover showed that search for dented-chin stimuli is governed by the same principles: The SA for dented-chin faces, too, only started to emerge when multiple items were visible at once, with the SA increasing in strength as the number of stimuli that was visible during a fixation increased. This indicates that the SA for dented faces is mediated by similar grouping processes that drive the SA for standard faces.

However, the evidence for the perceptual account admittedly depends on the assumption that the search asymmetries observed for standard and dented faces reflect the same underlying processing. Although Horstmann et al. (2010) found ratings of emotion similar for standard and dented faces, such ratings may not reflect the response properties of automatic and subconscious threat detectors. Hence, it could still be argued that the SA for standard faces is driven by emotional factors whereas the SA for dented faces is driven by purely perceptual factors. Experiment 3 critically tested this two systems hypothesis, by investigating whether and to what extent denting the faces may interfere with the processing of its emotional expression.

## Experiment 3

The aim of Experiment 3 was to test whether dented faces show the same distinctive processing characteristics as standard schematic faces, and in particular, whether the emotional expression is perceived as readily in dented faces as in standard faces. To that aim, we presented only a single emotional schematic face on each trial in Experiment 3 and assessed the readiness to perceive the emotional expression of dented versus standard faces by present-

ing the faces in their usual orientation or upside down. Importantly, the emotional expression of the faces was irrelevant to the task: Observers had to indicate whether the mouthline of the face was curved upward or downward with respect to *absolute standards*, while observers were instructed to ignore the relation of the mouthline to the face (that would allow discriminating happy from angry faces).

With respect to standard schematic faces, we would expect that observers still process the emotional expression of the faces, even when it is irrelevant to the task and the faces are presented upside down (e.g., Fallshore & Bartholow, 2003; Horstmann & Becker, 2008; Lipp, Price & Tellegen, 2009a; McKelvie, 1995; Valentine, 1988). If this is the case, then responses to inverted faces should produce costs, because the instructed response is incompatible with the (irrelevant) direction of the mouthline with respect to the face. For instance, in a friendly inverted face, the mouthline is curved downward according to absolute standards, but it is curved upward with respect to the face, so that the irrelevant feature is *incompatible* with the instructed response. By contrast, in upright faces, the mouthline is curved in the same direction according to absolute standards and in relation to the face, so that the irrelevant emotional expression is *compatible* with the instructed response. Hence, if emotional expression is processed automatically, then responses should be faster for response-compatible, upright faces than for response-incompatible, inverted faces (see Figure 1 for an overview of the conditions).

Moreover, comparing the degree of slowing between standard friendly and angry faces and the dented faces allows an evaluation of whether perceptual changes in the chin region of the dented faces indeed impaired automatic processing of their emotional expression. If the two-systems view is correct and the distortions to the chin-region indeed render the dented stimuli less face-like and hamper processing of their emotional expressions, then there should be no difference in RT to upright and inverted dented faces, because the compatibility effect critically depends on automatic processing of the emotional expression of faces. By contrast, if the emotional expression of dented faces is processed in a similar manner as that of the standard faces, then RT to inverted dented faces should be similarly slowed as RT to standard inverted faces.

To control for possible differences in detection of the emotional expression that is solely attributable to Gestalt differences between standard and dented faces or between upright and inverted faces, we included scrambled faces as control stimuli. The scrambled faces consisted of the same lines that made up the standard schematic faces; however, the dots representing the eyes were placed above and below the mouthline, so that the stimuli did not resemble a face anymore. Inversion effects that result from purely perceptual factors should also occur with the scrambled faces, whereas inversion effects resulting from the perception of the stimulus as a face (incompatibility effect) should be absent in the scrambled condition.

The incompatibility effect, which would in a first approximation be computed as difference between RT to inverted faces minus RT to upright faces (= inversion effect) can then be computed in a clearer fashion and excluding perceptual factors as “(RT to inverted face – RT to inverted scrambled face) – (RT to upright face – RT to upright scrambled face).”

Given that the previous experiments and Experiments 2A and 2B showed that the SA for angry standard faces and for friendly

dented faces is usually absent when only a single stimulus is visible, we would expect no differences between angry and friendly face targets in Experiment 3 (i.e., no SA).

## Method

**Participants.** Sixteen new subjects from the University of Queensland, Australia took part in the experiment as paid volunteers (\$10). All subjects had normal or corrected-to-normal vision and were naive as to the purpose of the experiment.

**Materials.** These were the same as in the previous experiments, with the exceptions that no eye tracker was used. Participants were seated in a dimly lit room, with their head resting in a chin rest, and viewed the screen from a distance of 57 cm.

**Stimuli.** Deviating from the previous experiments, the displays consisted of a single schematic face (diameter: 2.85 cm) that was presented either upright or inverted at the center of the display. Standard faces were created such that the eyes were 0.4 cm away from the upper facial outline, and the mouthline was 0.9 cm away from the lower facial outline. In scrambled faces, the mouthline was at exactly the same position as in the standard faces, but the eye-dots were positioned equidistantly above and below the mouthline, so that the lower dot was 0.4 cm away from the lower face outline (see Figure 1 for an example of the stimulus displays). Friendly and angry faces differed only in the direction of the mouthline (upward vs. downward curved, respectively). Standard faces had a round facial outline, whereas the facial outline of the dented stimuli was caved inward directly below the mouthline.

**Design.** The experiment consisted of the  $2 \times 2 \times 2 \times 2$  variables face-likeness (face vs. scrambled face), stimulus type (standard vs. dented face), emotional expression (friendly vs. angry), and stimulus orientation (upright vs. inverted). The variables face-likeness and stimulus type were varied across different blocks, with the order of blocks being counterbalanced across participants. The orientation of the face (upright/inverted) and its emotional expression (angry vs. friendly) was varied randomly within blocks, with the restriction that each face appeared equally often. Stimulus-response mappings were also counterbalanced. Participants completed 200 trials per block, resulting in 800 observations per participant.

**Procedure.** Each trial started with of a small black fixation cross ( $0.25 \text{ cm} \times 0.25 \text{ cm}$ ) at the center of the screen, which was presented for a time randomly chosen from 500 to 1,000 ms. Immediately afterward, the task-relevant stimulus was presented until the response, followed by a feedback display. The feedback consisted in the black printed words “Correct!” or “Wrong!” (in Arial Black, 13 pt.), which were presented centrally. Correct feedback was presented for 750 ms, wrong feedback remained on screen for 1,500 ms, to provide an extra incentive to avoid errors in this task. After an intertrial interval of 500 ms, in which a blank white screen was presented, the next trial started with the presentation of the fixation cross.

Before the experiment, participants were given written instructions about the task and stimuli in the next block. Participants were instructed to respond as fast as possible without making mistakes. On average, it took 40 minutes to complete the experiment.

## Results

**Data.** Excluding trials with RT shorter than 200ms or longer than 1,500 ms from all analysis resulted in a loss of 0.51% of the data. In the analysis of RT, only correct trials were included.

**Mean RT.** To examine the effect of face inversion, mean correct RT were first subjected to a  $2 \times 2 \times 2$  within-subjects ANOVA comprising the variables face-likeness (scrambled vs. nonscrambled), emotion (angry vs. friendly face), and orientation (upright vs. inverted). First, with respect to the *dented-chin faces*, the ANOVA showed significant main effects of all three variables (all  $p < .011$ , all  $\eta^2 > .35$ ), a marginally significant interaction between emotion and face-likeness [ $F(1, 15) = 4.5$ ;  $p = .051$ ;  $\eta^2 = .23$ ], and, more importantly, a significant interaction between orientation and face-likeness [ $F(1, 15) = 8.7$ ;  $p = .010$ ;  $\eta^2 = .37$ ]. As expected, RT to inverted stimuli were slower for face-like stimuli, where inversion rendered the mouthline-direction incom-

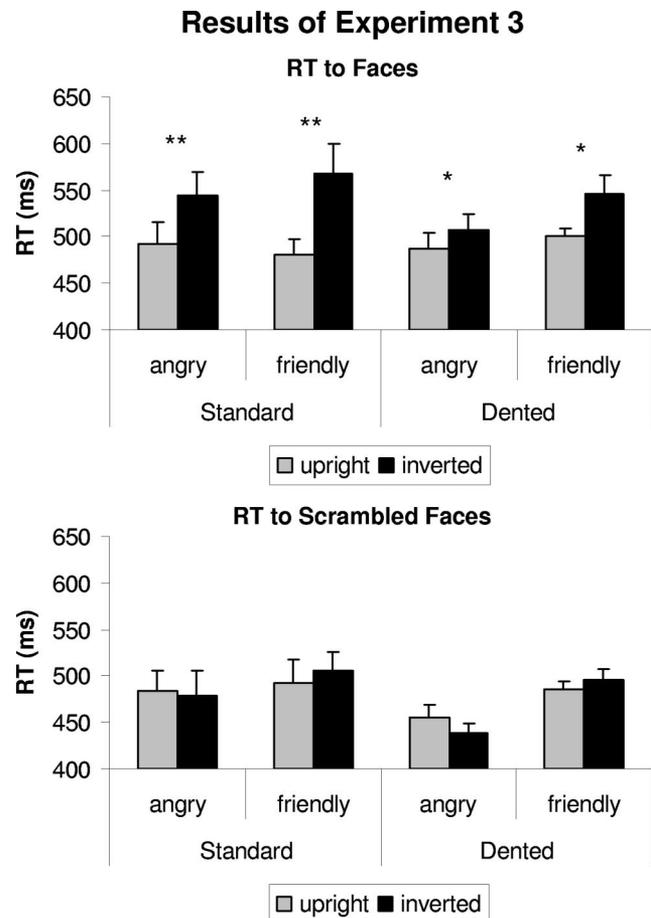


Figure 4. Mean RT in Experiment 3, where participants had to respond to the mouthline-direction of upright and inverted schematic faces that were presented individually. Responses to standard faces are on the left and responses to dented faces on the right of each figure, with the top panel depicting mean RT to faces, and the bottom panel depicting RT to scrambled faces. Elevated RT in the inverted condition indicate that the emotional expression of standard and dented faces interfered with the task, and thus, that it was processed despite the fact that it was irrelevant and the faces were presented upside down. \*  $p \leq .05$ . \*\*  $p \leq .01$ .

patible with the instructed response [ $F(1, 15) = 10.0$ ;  $p = .007$ ;  $\eta^2 = .40$ ], whereas orientation did not affect responses to scrambled faces ( $F < 1$ ; see Figure 4). There was a numerical trend for a stronger incompatibility effect with the friendly dented-chin face (44 ms) than for the angry dented-chin face (20 ms), but this trend was nonsignificant [ $F(1, 15) = 1.4$ ;  $p = .25$ ]. Moreover, the incompatibility effect was significant for both angry dented-chin faces [ $F(1, 15) = 4.5$ ;  $p = .05$ ;  $\eta^2 = .23$ ] and friendly dented-chin faces [ $F(1, 15) = 5.9$ ;  $p = .028$ ;  $\eta^2 = .28$ ].

The same analysis computed over the mean RT in responding to standard faces showed a significant main effect of face inversion [ $F(1, 15) = 17.5$ ;  $p = .001$ ;  $\eta^2 = .54$ ] and a significant interaction between emotion and face-likeness [ $F(1, 15) = 20.0$ ;  $p < .001$ ;  $\eta^2 = .57$ ]. The interaction was attributable to the fact that the inversion effect was significant only for the friendly and angry faces, where inverting the face rendered the mouthline incompatible with the instructed response [ $F(1, 15) = 20.2$ ;  $p < .001$ ;  $\eta^2 = .57$ ] but not for the scrambled faces ( $F < 1$ ). There was a numerical but nonsignificant trend for stronger incompatibility effects for the standard friendly face (52 ms) than for the angry face [87 ms;  $F(1, 15) = 1.8$ ;  $p = .20$ ], and incompatibility effects were significant for both angry [ $F(1, 15) = 9.9$ ;  $p = .009$ ;  $\eta^2 = .40$ ] and friendly faces [ $F(1, 15) = 14.0$ ;  $p = .002$ ;  $\eta^2 = .48$ ]. The fact that inversion costs occurred only with face-like stimuli but not scrambled faces demonstrates that inversion costs were not attributable to differences in the perception of single features, but to response incompatibility effects that arose from processing the direction of the mouthline relative to the face.

To assess whether the incompatibility effect was stronger with the standard faces than with the dented-chin faces, the inversion effects of the face-like stimuli were compared with one another. In a first analysis, only the face-like stimuli were included whereas the scrambled faces were disregarded. A  $2 \times 2 \times 2$  ANOVA with the variables stimulus type (standard vs. dented-chin face), emotion (friendly vs. angry face), and orientation (upright vs. inverted) showed significant main effects of emotion [ $F(1, 15) = 5.6$ ;  $p = .032$ ;  $\eta^2 = .27$ ] and orientation [ $F(1, 15) = 19.3$ ;  $p = .001$ ;  $\eta^2 = .56$ ]. The interaction between face type and orientation [ $F(1, 15) = 9.5$ ;  $p = .008$ ;  $\eta^2 = .39$ ] was also significant, reflecting that the incompatibility effect was stronger for the standard faces than for the dented-chin faces. The same results were obtained when, instead of the mean RT, the difference scores of RT to the standard faces – RT to the scrambled faces were entered into the ANOVA [main effect of emotion:  $F(1, 15) = 4.5$ ;  $p = .051$ ;  $\eta^2 = .23$ ; main effect of orientation:  $F(1, 15) = 18.8$ ;

$p = .001$ ;  $\eta^2 = .56$ ; face type  $\times$  orientation interaction:  $F(1, 15) = 5.3$ ;  $p = .036$ ;  $\eta^2 = .26$ ].

**Mean errors.** The mean error scores are depicted in Table 3 and were subjected to the same analyses. Analyzing the effects of face-likeness, emotion, and inversion first for the dented-chin stimuli yielded a significant main effect of face-likeness only [ $F(1, 15) = 15.9$ ;  $p = .001$ ;  $\eta^2 = .51$ ], reflecting that accuracy was higher for the scrambled faces than for the standard faces.

The same analysis computed over responses to the standard faces showed significant main effects for face-likeness [ $F(1, 15) = 10.2$ ;  $p = .006$ ;  $\eta^2 = .41$ ], emotion [ $F(1, 15) = 6.0$ ;  $p = .027$ ;  $\eta^2 = .29$ ], and inversion [ $F(1, 15) = 9.8$ ;  $p = .007$ ;  $\eta^2 = .40$ ]. Moreover, the interaction between face-likeness and orientation was significant [ $F(1, 15) = 8.6$ ;  $p = .010$ ;  $\eta^2 = .37$ ], reflecting that accuracy was significantly lower for inverted faces than for upright faces [ $F(1, 15) = 13.3$ ;  $p = .002$ ;  $\eta^2 = .47$ ], whereas the orientation had no effects on responses to scrambled faces ( $F < 1$ ). The inversion effect for face targets was in the correct direction, with more errors for inverted faces than for upright faces, indicating that the results were not attributable to a speed–accuracy trade-off.

**Discussion**

Consistent with earlier work and the results of Experiments 2A and 2B, the present results showed no SA for angry schematic faces in the single-item presentation condition of Experiment 3 (see Figure 4). As outlined above, this outcome is expected on various accounts, including the grouping account, and reinforces the view that the SA critically depends on the presence of multiple stimuli.

More importantly, the results of Experiment 3 showed that the emotional expression of face-like stimuli significantly interfered with the task, both for dented and standard schematic faces, while scrambled faces showed no interference effects. This demonstrates that both standard and dented faces were indeed perceived as faces with an emotional expression, which interfered with the task despite being task-irrelevant. Although the incompatibility effect was significantly stronger for the standard faces than for the dented faces, emotional expression significantly interfered with the task in both dented and standard faces. This finding invalidates the two-systems explanation, that the SA for dented faces is based on perceptual factors, because distortions to the chin region prevent processing of the emotional expression of dented faces. The sig-

Table 3  
Experiment 3: Mean Percent of Errors and the Standard Error of the Mean (in Parentheses)

	Standard stimuli				Dented stimuli			
	Faces		Scrambled faces		Faces		Scrambled faces	
	Angry target	Friendly target	Angry target	Friendly target	Angry target	Friendly target	Angry target	Friendly target
Upright	3.0% (0.9)	4.6% (1.0)	2.7% (0.7)	6.1% (1.8)	4.1% (1.1)	6.0% (1.8)	2.6% (0.6)	4.9% (1.4)
Inverted	7.4% (2.1)	8.9% (1.6)	2.3% (0.7)	6.9% (1.5)	5.0% (1.9)	6.4% (1.4)	1.5% (0.9)	4.9% (1.3)
Inversion effect	4.38	4.22**	-0.34	0.81	0.87	0.36	-1.12	-0.01

Note. Mean error scores of Experiment 3, where observers responded to the absolute mouthline-direction of a centrally presented single face.  
\*\*  $p < .01$ .

nificant reduction in interference costs for dented faces is more likely to be attributable to perceptual differences rather than emotional differences between the faces. Note that there was a trend for an opposite inversion effect in the scrambled angry dented face, as reflected in faster responses to inverted than to upright faces (see bottom panel of Figure 4). This trend would have reduced the incompatibility effect for the dented angry face when compared with the standard angry face.

In sum, the results support the claim of the perceptual account, that dented and standard schematic are both perceived as faces with a particular emotional expression, and that the SA for angry standard faces is not driven by emotion but perceptual saliency.

Experiment 3 showed that the emotional expression of dented faces is processed even when it is task-irrelevant and when the faces are presented upside down. This would seem to provide strong support for the perceptual account, because, first, there is no reason to believe that the emotional expression of dented faces is inaccessible in visual search tasks, where only upright faces are presented and observers are instructed to actively search for an angry or friendly face. Second, the results indicate that dented faces share some of the processing characteristics that apply specifically to faces and distinguish emotional schematic faces from other complex patterned stimuli. Previous research has shown that processing of emotional faces differs from processing of other stimuli in that it is (1) automatic (e.g., Horstmann & Becker, 2008; Lipp, Price & Tellegen, 2009a,b); (2) holistic, insofar as the elements of the face cannot be easily decomposed into component features (e.g., Farah et al., 1998), and (3) at least partially orientation-invariant; that is, whereas the processing of facial identity is greatly impaired when the face is presented upside down, inverted friendly and angry faces can still be recognized as faces with a particular emotional expression, indicating that processing of emotional content remains largely intact even when the face is presented in an unusual orientation (e.g., Fallshore & Bartholow, 2003; McKelvie, 1995; Valentine, 1988). The present results indicate that dented emotional faces show all of these processing characteristics as well, indicating that denting the faces did not diminish the face-likeness or automatic processing of their emotional expression.

To note, the automaticity of processing of emotional expressions is supported only in a limited sense, because observers in Experiment 3 were instructed to attend to the face at fixation. Hence, the results can only show that facial expressions are processed automatically when observers attend to the face. However, this does not affect our conclusions. As Bargh (1989, 1992) noted, there are hardly any effects that fulfill the classical criteria of automaticity as involving only involuntary, unintentional, autonomous, unconscious, and effortless processing. The majority of automatic processes are contingent on a previous decision to attend to the stimulus. For example, the Stroop effect and feature priming effects are eliminated when observers do not actively attend to the stimulus (e.g., Becker, 2007), and there is evidence that the same may be true for emotional faces (e.g., Pessoa, Padmala & Morland, 2005; Moser, Hajcak, Bukay, & Simons, 2006). Previous results and the present findings are consistent with Bargh's notion of *conditional automaticity*, that facial expressions are processed automatically given a prior decision to attend to faces, and in the following, we will use the term "automatic" to refer to this notion of conditional automaticity.

More importantly, the possibility that emotional content is *not* available preattentively cannot harm our conclusions: In this case, emotional expressions could not guide attention and hence, the SA for angry faces could not be attributable to emotional factors, consistent with the perceptual account.

## General Discussion

The present results provide the perhaps best evidence available today for three important claims: (1) that the SA for angry schematic faces is based on perceptual and not emotional factors, (2) that the SA is mainly driven by grouping of the distractors (and not target capture or de-allocation of attention from selected distractors), and (3) that processing of emotional expressions is to some extent automatic and orientation-invariant. We discuss each of these in turn below.

### The SA Is Driven by Perceptual and Not Emotional Factors

In the past, the perceptual account of the SA for angry schematic faces was mainly supported by two observations: First, Coelho, Cloete, and Wallis (2009) tested a large range of non-face-like stimuli which preserved the salient perceptual characteristics of angry versus friendly schematic faces (i.e., eyebrows and mouthline-curvature) and showed that the SA could also be obtained with these non-face-like stimuli. Second, Horstmann et al. (2010) showed that the SA could be reversed by changing the perceptual characteristics of the stimuli that left their emotional expression intact (i.e., creating dented faces; e.g., Horstmann et al., 2010). These findings are difficult to explain on an emotional account of the SA, which attributes the SA to differences in the emotional expression of schematic faces.

Proponents of the emotional factor view could however argue that these findings cannot refute the emotional factors explanation of the SA for angry schematic faces: The emotional factor explanation does not make any predictions about non-face-like stimuli, so the results of Coelho and colleagues (2009) are, strictly speaking, not inconsistent with the emotional factor account. Similarly, it can be argued that the alterations to the facial outline of the stimuli in Horstmann et al.'s (2010) study may have destroyed effects of the emotional expression, so that the SA for dented faces must be regarded as a new SA different from the SA for standard schematic faces, which could still be attributable to emotional factors. Experiments 1 and 2 probed into this possibility by measuring eye movement behavior in free viewing versus restricted viewing conditions and found that the results pattern was exactly mirror-reversed for dented faces, showing no evidence that the SAs are based on different search mechanisms.

However, the possibility that the respective SAs are driven by different factors (emotional in the standard case, perceptual in the dented faces) is strictly speaking impossible to reject experimentally, because differences in the underlying processes do not have to translate into differences in behavioral measures. On the other hand, the two factors explanation critically depends on the assumption that standard faces and dented faces have different emotional characteristics: In particular, proponents of the emotional factors view have to deny that denting the emotional faces changed only the perceptual characteristics of the faces while leaving their

emotional expressions intact, as was claimed by Horstmann et al. (2010). Experiment 3 was designed to critically test this assumption and showed that the emotional expression of dented faces is processed even when (1) emotional expression is task-irrelevant and processing of it interferes with the task, and (2) when the face is presented upside down. These results indicate that denting the faces did not eradicate immediate and automatic processing of their emotional expression, contrary to the two systems explanation. Taken together, the combined findings from the present study render it very improbable that the SA for angry schematic faces is driven by emotional factors.

Similar doubts about the emotional factor explanation have been raised with regard to photographic images: Whereas more efficient search for photorealistic angry faces than friendly faces was initially attributed to emotional factors ("anger superiority effect; e.g., Hansen & Hansen, 1988), later studies showed that efficient detection of angry faces was due to an artifact of stimulus editing that left dark smudges on the angry face. When the dark regions were removed, the SA for the photorealistic faces was reversed, with faster search for a friendly face among angry faces than vice versa (Purcell, Stewart, & Skov, 1996), and this finding has since then been replicated in a number of studies (e.g., Lipp, Price, & Tellegen, 2009a, 2009b; Juth, Lundqvist, Karlsson, & Öhman, 2005). The widely inconsistent results have led many researchers to propose that the SAs with photographic images are mostly determined by perceptual factors. In different stimulus sets, search efficiency has been shown to depend on salient features present in the mouth region (e.g., Horstmann & Bauland, 2006) or in the eye region (Fox & Damjanovic, 2006), and it has been hypothesized that the visibility of teeth may be important in driving the SA (e.g., Lipp et al., 2009a). Thus, it may seem plausible to apply the perceptual grouping account to explain SAs found with photographic faces. However, with regard to photorealistic faces, the quest for a *single* perceptual property or face region driving the SA may well prove futile: In such complex stimulus sets, angry faces usually differ in more than one feature from friendly faces, and it cannot be expected that all observers will adopt the same search strategy (i.e., in attending to the mouth rather than to the eye region).

Could SAs observed with other fear-relevant stimuli such as snakes and spiders also be attributable to perceptual rather than emotional factors? On the one hand, it is plausible that pictures of fear-relevant stimuli may have been perceptually more salient than the control stimuli (e.g., mushrooms, flowers, fish, or birds; e.g., e.g., Eastwood et al., 1995; Rinck et al., 2005; Öhman, Flykt & Estevez, 2001), for instance, if snakes and spiders were more homogeneously colored and/or presented against a more homogeneous background. This would create a stronger figure-ground contrast, so that fear-relevant stimuli may have similarly hampered grouping or caused interference on a perceptual level when they constituted the distractors. However, the perceptual factor explanation has difficulty explaining the results from some clinical studies; in particular, the finding that observers who are anxious of a particular group of potentially threatening stimuli are especially slow to search through them when they constitute the distractors (Lipp, Waters, Derakshan, & Logies, 2004; Lipp & Waters, 2007; Öhman et al., 2001; but see Becker & Rink, 2006; Miltner, Krieschel, Hecht, Trippe, & Weiss, 2004). These results indicate that fear-provoking stimuli may indeed actively divert attention and/or produce difficulties in disengaging attention (but see Lipp et al., 2004).<sup>2</sup>

In conclusion, perceptual factors explanations such as the perceptual grouping account have proven useful in explaining some of the SAs previously attributed to emotional factors, but the current state of evidence renders it unlikely that all SAs are completely determined by perceptual factors.

### Perceptual Grouping Accounts for the SA for Angry Schematic Faces

The present experiments provide the perhaps most direct evidence for a grouping account of the SA for angry schematic faces. Experiments 1 and 2 clearly showed that the SA is attributable to distractor-related effects rather than target-related effects (e.g., target capture), and specifically, the fact that friendly faces can be grouped more easily than angry faces when they constitute the distractors. These findings are in line with earlier studies (e.g., Horstmann & Bauland, 2006; Horstmann et al., 2010) and reinforce the view that grouping plays an important role for determining search efficiency: Experiment 2 showed that search performance critically depended on the size of the gaze-contingent window, supporting the view that the major factor determining search efficiency is the size of the attentional window, which in turn determines the number of items that can be rejected in parallel (Zelinsky & Sheinberg, 1997; Williams, Reingold, Moscovitch, & Behrmann, 1997).

The finding that the SA is driven by grouping which is in turn based on perceptual factors is an important finding, especially because it can easily be imagined that perceptual grouping also accounts for SAs observed with other stimuli.

Asymmetrical search is ubiquitous and has been found to occur with stimulus pairs across all stimulus dimensions. For instance, in the orientation dimension it has been found that search for a tilted bar among vertical bars is faster than vice versa (e.g., Treisman & Gormican, 1988; Wolfe, 2001); with shapes, search for a Q among O was found faster than the reverse (e.g., Treisman & Souther, 1985), and with moving stimuli, motion could be detected faster among stationary items than vice versa (e.g., Wolfe, 2001). From early on, the presence of asymmetrical search in seemingly symmetrical conditions has been a mystery, and it still presents a challenge for current theories of visual search: Note that differences in search performance that occur when the target and distractors are swapped cannot be explained by factors that are known to modulate search efficiency, as for instance, feature contrast, target-distractor similarity, or distractor-distractor dissimilarity (e.g., Duncan & Humphreys, 1989), because swapping targets and distractors does not change any of these factors. Correspondingly, the

<sup>2</sup> One complicating factor is that threat detectors could be feature detectors themselves that respond, for instance, to the wavy shape of a snake or the eight-legged shape of a spider (e.g., Öhman & Mineka, 2001). Such an account cannot be distinguished behaviorally from a perceptual factors explanation. Some fMRI studies suggest that potentially threatening stimuli can activate the amygdala automatically and without deploying attention to the stimulus, which has been taken as evidence for threat detectors (Carlsson et al., 2004 and Straube, Mentzel, & Miltner, 2006; but see Alpers et al., 2009). However, so far, there is no evidence that activity in fear-relevant modules modulates search performance, which would be needed to demonstrate the existence of threat detectors (i.e., the ability of fear-relevant modules to improve detection).

existence of SAs has mostly been explained by preferences that are hard-wired into the visual system (e.g., Treisman & Souther, 1985; Wolfe, 1998, 2001; for a noteworthy exception see Rauschenberger & Yantis, 2006).

By contrast, the perceptual grouping account offers an explanation for the SA for angry schematic faces that is not based on hard-wired and thus inexplicable factors but instead draws on factors that are already known to modulate search efficiency. According to the perceptual grouping account, grouping is facilitated for friendly schematic faces because they have a simpler Gestalt, which is attributable to the fact that the mouth-chin region is organized in parallel lines which create less noise or less of a feature contrast than in the angry face, where the mouthline is curved perpendicular to the facial outline. The hypothesis that the SA is driven by these differences in the feature contrast or Gestalt of the faces was tested and confirmed by showing that the SA could be reversed by denting the chin-region of the stimuli (which forced the angry faces into the simpler Gestalt with the parallel chin-mouth set-up). The perceptual account arrived at this prediction by assuming that the attention-driving capacity of the critical and task-relevant feature (i.e., mouthline direction) has to be assessed relative to a nominally irrelevant part of the stimulus (i.e., facial outline), so that the context or background of the task-relevant feature has to be taken into account in saliency computations.

In this respect, it is noteworthy that a similar hypothesis has also been proposed to account for SAs in simpler search tasks. Rosenholtz (1999) argued that a moving target can be found faster among stationary distractors than vice versa because the whole background is stationary (e.g., the background of the display, but also the monitor frame and other visible items). This creates a larger feature contrast for the moving stimulus than for the stationary stimulus, which renders the moving stimulus more salient and facilitates detection. Similarly, it has been argued that the SA for a tilted item among vertical distractors is attributable to the fact that the vertical items are aligned with the monitor frame, which renders them less salient than the tilted items which differ more strongly from the background. In line with this view, it has also been shown that the SA for tilted lines can be reversed when the monitor itself is tilted, so that the tilted target is then aligned with the sides of the monitor frame (Treisman & Gormican, 1988; see also May & Zhaoping, 2009; Treisman & Souther, 1985; Wolfe, 1998, 2001).

Li (1999, 2002) recently proposed that such and similar SAs may be a result of an iso-orientation suppression mechanism that reduces pop-out for stimuli with a similar orientation as the context (via horizontal intracortical connections in V1; see also Nothdurft, Gallant, & van Essen, 2000; Wolfe, 1998). Such an iso-orientation suppression mechanism may also account for the present findings, as iso-feature suppression would have reduced pop-out of the stimuli where the mouthline conformed to the lower face outline (i.e., standard friendly and dented angry stimuli; see also May & Zhaoping, 2009).

The perceptual grouping account extends on these findings and the corresponding explanations by extending the scope to more complex stimuli that do not pop out, but typically produce inefficient or semiefficient search (or a “relative SA”; see Horstmann & Becker, 2008), and demonstrating that, in inefficient or semiefficient search, more or less efficient grouping of the distractors is primarily responsible for asymmetrical search.

## Faces Are Processed Automatically and to Some Extent Orientation-Invariant

Previous research has shown that faces are processed holistically: Elements of a face are not perceived as individual features that have to be combined to form the impression of face, but rather individual features are already perceived in virtue of their role of being part of a face (e.g., Farah et al., 1998). Evidence for this hypothesis can be derived from studies showing that holistic processing of faces occurs even when it is not required: Embedding a facial feature (e.g., nose or mouth) in a face context usually hampers feature-specific processing, indicating that holistic processing is to some extent automatic and can interfere with feature-specific processing. The results of Experiment 3 extend on this research by showing that processing of the emotional expression of faces (happy, angry) proceeds automatically, as expressions were processed despite the fact that they were completely irrelevant to the task and harmed performance. This suggests that the processing of emotional expression in an integral part of face processing, and proceeds automatically—at least when observers are instructed to attend to the face (see also Horstmann & Becker, 2008).

More importantly, Experiment 3 showed that the face elements were assessed in relation to the whole face despite the fact that the faces were presented upside down in the critical condition. This indicates that automatic processing of face characteristics such as the emotional expression does not critically depend on the orientation of the face. These results are consistent with earlier studies that showed that the emotional expression of faces can still be recognized reasonably accurately even when faces are presented upside down: Although performance in emotion recognition tasks was usually impaired when the faces were presented upside down, accuracy was well above chance level in all studies (e.g., Fallshore & Bartholow, 2003; McKelvie, 1995).

In previous studies, corresponding results were often interpreted to show that processing of emotional expressions is feature-specific and *not* holistic. This is in accord with traditions, as holistic processing has usually been inferred whenever inverted presentation conditions led to performance decrements (whereas feature-specific processing modes have been assumed when inverted presentation conditions did not lead to performance decrements; e.g., Valentine, 1998). However, note that both the stimulus materials and the task of previous studies differed from the present task: In previous studies, observers were usually asked to identify or recognize faces from photographic images that they had previously learnt to distinguish from one another. In this identification or recognition tasks, inverting the photographic images usually greatly impaired performance (e.g., Valentine, 1998). This indicates that faces are not identified via specific features; for instance, a person is not identified by relying on a specific shade of green of the eyes, or by committing a specific size of the lips to memory, which would have been accessible in both upright and inverted conditions. Instead, face recognition apparently relies on assessing subtle relations between different face elements, which is a hallmark of holistic or configural processing.

Studies showing that processing of the emotional expression of faces is unimpaired by face inversion have consequently concluded that identification of the emotional expression does not require holistic processing but is based on feature-specific processing (e.g., Lipp et al., 2009a, 2009b; Fallshore & Bartholow, 2003;

McKelvie, 1995). However, this interpretation does not seem applicable to the schematic faces used in Experiment 3: Whereas emotional expressions in photographic images could in principle be identified or recognized by a salient feature, such as the visibility of the teeth in the smiling face, the ability to distinguish between emotional expressions in inverted schematic faces cannot be explained by such feature-specific processing. In the present study and other studies testing schematic faces, happy and angry faces differ only in the direction of the mouthline. Moreover, the inverted angry face has exactly the same mouthline as the happy upright face, and happy inverted faces have exactly the same mouthline as angry inverted faces. Hence, feature-specific processing of the mouth alone cannot discriminate between happy and angry expressions. Consequently, the ability to process emotional expressions in inverted schematic faces cannot be interpreted as evidence for nonholistic, feature-specific processing. Instead, discriminating between different emotional expressions requires assessing the critical face element in relation to the face. This implies that a mental representation of the face existed, which means that the face was processed holistically despite the fact that it was inverted.

The results of Experiment 3 highlight the fact that inverted faces can be processed holistically, because the mouthline direction was evidently assessed relative to the face although observers were asked to respond to the absolute mouthline direction. This indicates that the faces were processed holistically despite the fact that (1) the task required only feature-specific processing of the mouth and (2) the faces were inverted. Hence, Experiment 3 demonstrates that emotional expressions in inverted schematic faces are processed both holistically and in an automatic fashion (in the sense of conditional automaticity; Bargh, 1992).

The conclusion that processing of inverted faces can proceed holistically can only be averted if it were possible that the mouthline direction was processed in relation to the face without simultaneously processing the emotional expression of the face. However, it is difficult to see how observers could know that the mouthline is curved upward with respect to the face without knowing that the face is smiling. Clinical research has shown some extraordinary dissociations in face processing (e.g., Calder & Young, 1995; Palermo & Rhodes, 2007); however, so far, there have been no reports of neurological conditions that would result in impairments in recognizing emotional expressions that left abilities to judge the orientation of the mouthline of inverted faces intact. Without further evidence to the contrary, we should therefore conclude that the results of Experiment 3 indeed demonstrate holistic processing of inverted faces.

The fact that inverting a face impairs identification of faces to a larger extent than recognition of friendly and angry emotional expressions could be attributable to the fact that the two tasks rely on different processes. In line with this view, a multitude of clinical and neurophysiological studies indicate that processes of identification of particular faces and recognition of emotional expression recruit different brain areas (for a review, see Calder & Young, 1995). For instance, stroke patients have occasionally been reported to show deficits in face recognition without an accompanying deficit in recognizing emotional expressions, and the reverse has also been reported (e.g., Palermo & Rhodes, 2007; see also Bar, Neta, & Linz, 2006).

However, Calder and Young (1995) argued that the available evidence for selective impairments in emotion recognition is quite weak and could be attributable to experimental artifacts. Selective impairments in face identification do not necessarily signify important differences in the underlying processes but could be attributable to the fact that face identification is more difficult than recognizing emotional expressions; hence, face identification may require additional processes that are selectively impaired in patients, leading to the dissociation in performance. According to this view, face identification and identification of the emotional expression could be quite closely related, with the magnitude of inversion effects depending on the difficulty of the task (e.g., Calder & Young, 2005).

The latter view, apart from being more parsimonious, also seems to be quite plausible, when we consider that recognizing the emotional expression of a face requires only information about the relation of a salient feature (e.g., the mouth) to the face, whereas face identification will often require more fine-grained and subtle distinctions. Although it seems plausible to assume that face processing must be holistic when inversion impairs performance, testing inverted faces probably does not provide an ideal test of holistic or configural processing. Note that inverting a complex stimulus will also distort individual features: Whereas features such as color, luminance and size are presumably unaffected by inversion, the shape and orientation of single features (e.g., slant of the eyes and eye brows, shape of the nose and mouth) are certainly altered in inverted faces. Hence, as long as it is unknown whether and to what extent observers rely on such orientation-dependent feature-specific information in a particular task, impaired performance upon inversion cannot be clearly attributed to holistic or configural processing (see also Valentine, 1988). This is all the more so, because the present Experiment 3 showed that inverting the face apparently left some holistic information intact or at least readily accessible to the observer. Clearly, further research is necessary to assess whether and to what extent face inversion experiments can substantiate conclusions about how the faces are processed. What seems to be clear from the present study is that processing of the emotional expression of schematic faces is possible even when they are inverted, whereby holistic processing proceeds to some extent automatically.

## References

- Alpers, G. W., Gerdes, A. B. M., Lagarie, B., Tabbert, K., Vaitl, D., & Stark, R. (2009). Attention and amygdala activity: An fMRI study with spider pictures in spider phobia. *Journal of Neural Transmission*, *116*, 747–757.
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*, 269–278.
- Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. A. Bargh & J. S. Uleman (Eds.), *Unintended thought* (pp. 3–51). New York: Guilford Press.
- Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology*, *105*, 181–199.
- Becker, E. S., & Rinck, M. (2006). Sensitivity and response bias in fear for spiders. *Cognition & Emotion*, *18*, 961–976.
- Becker, S. I. (2007). Irrelevant singletons in pop-out search: Attentional capture or filtering costs? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 764–787.

- Becker, S. I. (2008a). Can intertrial effects of features and dimensions be explained by a single theory? *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1417–1440.
- Becker, S. I. (2008b). The stage of priming: Are intertrial repetition effects attentional or decisional? *Vision Research*, *48*, 664–684.
- Becker, S. I. (2010). The role of target-distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology: General*, *139*, 247–265.
- Becker, S. I., & Horstmann, G. (2009). A feature weighting account of priming in conjunction search. *Attention, Perception, & Psychophysics*, *71*, 258–272.
- Calder, A. J., & Young, A. W. (1995). Understanding the recognition of facial identity and facial expression. *Nature Reviews and Neuroscience*, *6*, 641–651.
- Carlsson, K., Petersson, K. M., Lundqvist, D., Karlsson, A., Ingvar, M., & Öhman, A. (2004). Fear and the amygdala: Manipulation of awareness generates differential cerebral responses to phobic and fear-relevant (but nonfeared) stimuli. *Emotion*, *4*, 340–353.
- Coelho, C. M., Cloete, S., & Wallis, G. (2009). The face-in-the-crowd effect: When faces are just cross(es). *Journal of Vision*, *10*, 1–14.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827–1837.
- 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 & Psychophysics*, *63*, 1004–1013.
- Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception & Psychophysics*, *65*, 352–358.
- Eastwood, J. D., Smilek, D., Oakman, J., Farvolden, P., Van Ameringen, M., Mancini, C., & Merikle, P. M. (1995). Individuals with social phobia are biased to become aware of negative faces. *Visual Cognition*, *12*, 159–179.
- Fallshore, M., & Bartholow, J. (2003). Recognition of emotion from inverted schematic face drawings. *Perception and Motor Skills*, *96*, 236–244.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychological Review*, *105*, 482–498.
- Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research*, *37*, 617–631.
- 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. J., 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., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition and Emotion*, *16*, 355–379.
- 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.
- Horstmann, G. (2007). Preattentive face processing: What do we learn from visual search experiments? *Visual Cognition*, *15*, 799–833.
- Horstmann, G. (2009). Visual search for affective faces: Stability and variability of search slopes with different instances? *Cognition and Emotion*, *23*, 355–379.
- Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion*, *6*, 193–207.
- Horstmann, G., & Becker, S. I. (2008). Attentional effects of negative faces: Top-down contingent or involuntary? *Perception & Psychophysics*, *70*, 1416–1434.
- Horstmann, G., Becker, S. I., Bergmann, S., & Burghaus, L. (2010). A reversal of the search asymmetry favoring negative schematic faces. *Visual Cognition*, *18*, 981–1016.
- Horstmann, G., Heumann, M., & Borgstedt, K. (2006a). Flanker effects with faces may depend on perceptual as well as emotional differences. *Emotion*, *6*, 28–39.
- Horstmann, G., Scharlau, I., & Ansorge (2006b). More efficient rejection of happy than angry face distractors in visual search. *Psychonomic Bulletin & Review*, *13*, 1067–1073.
- 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.
- Li, Z. (1999). Contextual influences in V1 as a basis for pop out and search asymmetry in visual search. *Proceedings of the National Academies of Science USA*, *96*, 10530–10535.
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Science*, *6*, 9–16.
- Lipp, O. V., & Derakshan, N. (2005). Attentional bias to pictures of fear-relevant animals in a dot probe task. *Emotion*, *5*, 365–369.
- Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009a). Emotional faces in neutral crowds: Detecting displays of anger, happiness, and sadness on schematic and photographic images of faces. *Emotion*, *33*, 240–260.
- Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009b). No effect of inversion on attentional and affective processing of facial expressions. *Emotion*, *9*, 248–259.
- Lipp, O. V., & Waters, A. M. (2007). When danger lurks in the background: Attentional capture by animal fear-relevant distractors is specific and selectively enhanced by animal fear. *Emotion*, *7*, 192–200.
- Lipp, O. V., Waters, A. M., Derakshan, N., & Logies, S. (2004). Snakes and cats in the flower bed: Fast detection is not specific to pictures of fear-relevant animals. *Emotion*, *4*, 233–250.
- Lundqvist, D., & Öhman, A. (2005). Emotion regulates attention: The relation between facial configurations, facial emotion, and visual attention. *Visual Cognition*, *12*, 51–84.
- May, K. A., & Zhaoping, L. (2009). Effects of surrounding frame on visual search for vertical or tilted bars. *Journal of Vision*, *29*, 1–19.
- McKelvie, S. I. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology*, *34*, 325–334.
- Miltner, W. H. R., Krieschel, S., Hecht, H., Trippe, R., & Weiss, T. (2004). Eye movement and behavioral responses to threatening and nonthreatening stimuli during visual search in phobic and nonphobic subjects. *Emotion*, *4*, 323–339.
- Moser, J. S., Hajcak, G., Bukay, E., & Simons, R. F. (2006). Intentional modulation of emotional responding to unpleasant pictures: An ERP study. *Psychophysiology*, *43*, 292–296.
- Nothdurft, H. C. (1993). Faces and facial expressions do not pop out. *Perception*, *22*, 1287–1298.
- Nothdurft, H. C., Gallant, J. L., & van Essen, D. C. (2000). Response profiles to texture border patterns in area V1. *Visual Neuroscience*, *17*, 421–436.
- Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*, 466–478.
- 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). Fear, phobias, and preparedness: Towards and evolved module of fear and fear learning. *Psychological Review*, *108*, 483–522.
- Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neurophysiologica*, *45*, 75–92.
- Pessoa, L., Padmala, S., & Morland, T. (2005). Fate of unattended fearful

- faces in the amygdala is determined by both attentional resources and cognitive modulation. *NeuroImage*, 28, 249–255.
- 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.
- Rauschenberger, R., & Yantis, S. (2006). Perceptual encoding efficiency in visual search. *Journal of Experimental Psychology: General*, 135, 116–131.
- Rinck, M., Reinecke, A., Ellwart, T., Heuer, K., & Becker, E. S. (2005). Speeded detection and increased distraction in fear of spiders: Evidence from eye movements. *Journal of Abnormal Psychology*, 114, 235–248.
- Rosenholtz, R. (1999). A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, 39, 3157–3163.
- Straube, T., Mentzel, H. J., & Miltner, W. H. (2006). Neural mechanisms of automatic and direct processing of phobogenic stimuli in specific phobia. *Biological Psychiatry*, 59, 162–170.
- Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). On the time course of top-down and bottom-up control of visual attention (pp. 105–125). In S. Monsell & J. Driver (Eds.). *Attention & Performance* (Vol. 18). Cambridge: MIT Press.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15–48.
- Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, 114, 285–310.
- Valentine, T. (1988). Up-side down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79, 471–491.
- Williams, D. E., Reingold, E., Moscovitch, M., & Behrmann, M. (1997). Patterns of eye movements during parallel and serial visual search tasks. *Canadian Journal of Experimental Psychology*, 51, 151–164.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238.
- Wolfe, J. M. (1998). Visual search. In Pashler H. (Ed.), *Attention* (pp. 13–73). London, UK: UCL Press.
- Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception and Psychophysics*, 63, 381–389.
- Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 244–262.
- Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science*, 8, 448–453.

Received September 29, 2010

Revision received January 30, 2011

Accepted April 6, 2011 ■

### **Correction to Becker et al. (2011)**

The Online First version of the article “Perceptual Grouping, Not Emotion, Accounts for Search Asymmetries With Schematic Faces,” by Stefanie I. Becker, Gernot Horstmann, and Roger W. Remington (*Journal of Experimental Psychology: Human Perception and Performance*, 2011, Advance online publication. August 1, 2011. doi:10.1037/a0024665), had errors in Figure 2. All versions of this article have been corrected.

DOI: 10.1037/a0026026