Do Emotional Faces Capture Attention, and Does this Depend on Awareness? Evidence from the Visual Probe Paradigm.

Nicholas Hedger[a], Matthew Garner[b,c], Wendy J. Adams[b]

[a]Psychology, University of Reading, Reading, RG6 6AL, UK

[b] Centre for Vision and Cognition, Psychology, University of Southampton, Southampton, SO17 1BJ, UK

[c] Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, SO17 1BJ, UK

Corresponding author: Nicholas Hedger.

Email: N.Hedger@reading.ac.uk

Phone: +447742142858

Address: Psychology, Reading, Reading, RG6 6AL, UK.

Word Count: 8294

Abstract

The visual probe (VP) paradigm provides evidence that emotional stimuli attract attention. Such effects have been reported even when stimuli are presented outside of awareness. These findings have shaped the idea that humans possess a processing pathway that detects evolutionarily significant signals independently of awareness. Here, we addressed two unresolved questions: First, if emotional stimuli attract attention, is this driven by their affective content, or by low-level image properties (e.g. luminance contrast)? Second, does attentional capture occur under conditions of genuine unawareness? We found that observers preferentially allocated attention to emotional faces under aware viewing conditions. However, this effect was best explained by low-level stimulus properties, rather than emotional content. When stimuli were presented outside of awareness (via continuous flash suppression or masking), we found no evidence that attention was directed towards emotional face stimuli. Finally, observer's awareness of the stimuli (assessed by d prime) predicted attentional cuing. Our data challenge existing literature: First, we cast doubt on the notion of preferential attention to emotional stimuli in the absence of awareness. Second, we question whether effects revealed by the VP paradigm genuinely reflect emotion-sensitive processes, instead suggesting they can be more parsimoniously explained by low-level variability between stimuli.

Keywords: threat; emotion; attention; awareness; visual probe

Public Significance Statement

Emotionally salient stimuli (such as fearful faces) are prioritised in attention, even when they are presented outside of awareness. Moreover, such effects are often found to be larger in anxious populations, suggesting that emotion sensitive mechanisms that operate without awareness may be involved in the aetiology/ maintenance of anxiety disorders. However, the mechanisms underlying such 'emotional attention' effects remain unclear. Here we show that i) emotional stimuli only attract attention under conditions where observers are aware of stimuli. ii) preferential attention to emotional faces is best explained by low-level stimulus properties (e.g. luminance contrast) rather than emotion-sensitive processes. Our study highlights the need for careful experimental control in basic and clinical research investigating the link between emotion and attention.

Human visual perception has limited capacity and must direct resources towards salient stimuli, events and spatial locations. Many behavioural studies suggest that emotionally salient (particularly threatening) stimuli attract our attention (Armony & Dolan, 2002; Ohman, Flykt & Esteves, 2001; Vuilleumier & Schwartz, 2001). The visual probe paradigm provides evidence of this effect. On a typical trial, an emotionally salient and a neutral target stimulus are presented on either side of a central fixation cross, before a probe (usually a small dot or arrow) appears at the location preceded by either the emotional stimulus (valid trial) or neutral stimulus (invalid trial). Observers then make a speeded response to indicate the location or orientation of the probe (left vs. right, or pointing up or down). Responses are typically faster in valid trials than invalid trials, suggesting that spatial attention has been preferentially allocated to the location of the emotional stimulus (Bar-Haim, Lavy, Pergamin, Bakermans-Kranenburg, & van Ijzendoorn, 2007). There is tremendous interest in understanding the mechanisms of this selection process - how does the visual system prioritize stimuli that are most important to its survival?

Evolutionary theories suggest that humans possess an independent, sub-cortical visual pathway that operates without awareness and rapidly directs processing resources towards threatening stimuli (Garrido, 2012; Tamietto & de Gelder, 2010). This theory has intuitive appeal - it may take hundreds of milliseconds for retinal stimulation to generate a conscious percept (Koch, 2004; Sekar, Findley, Poeppel, & Llinás, 2013). If threats could modulate an observer's behaviour rapidly and independently of their conscious registration, survival odds would be increased (Morris, Öhman, & Dolan, 1999). This notion is intriguing, because it suggests that there are specialised ways of (and independent neural substrates for) prioritising affective stimuli. Moreover, such an idea has influenced thinking about clinical disorders. For instance, dysfunction in the systems involved with preconscious threat detection are thought to underlie the hypersensitivity to threat and maladaptive perceptual biases exhibited by individuals with anxiety disorders (Mogg & Bradley, 1998; Ohman & Mineka, 2001).

Evidence for the unconscious prioritisation of threat has typically relied on measuring responses to stimuli that are presented to observers outside of awareness (Kim & Blake, 2005). A long history of observations from paradigms such as backward masking, binocular rivalry and continuous flash suppression (CFS) has revealed that threat stimuli suppressed from awareness can nonetheless elicit adaptive changes in neural activity (Jiang & He, 2006; Whalen et al., 2004; Williams, Morris, McGlone, Abbott, & Mattingley, 2004) and physiological arousal (Lapate, Rokers, Li, & Davidson, 2013; Ohman & Soares, 1994).

Behaviourally, the masked visual probe (MVP) paradigm has provided evidence that threat stimuli receive prioritized processing in the absence of awareness. In a modification of the standard visual probe design, target stimuli are presented briefly, and then replaced with a masking pattern. The small stimulus onset asynchrony (SOA) between the target and mask (usually ~17 or ~33 milliseconds) means that observers typically report perceiving the mask, but not the preceding target stimulus (Wiens & Ohman, 2007): visual presentation of the target stimulus is dissociated from awareness of it. Thus, the MVP paradigm can be employed as a tool to examine attentional orienting to emotionally salient stimuli in the absence of their conscious registration.

In a recent meta-analysis of the MVP paradigm (Hedger, Gray, Garner, & Adams, 2016) we found that the magnitude of threat-related bias (i.e. the valid vs. invalid response time (RT) difference) across all stimulus types (including fear and angry faces, negative words and images from the International Affective Picture System) tends to be small (Cohen's $d_z = 0.28$). Our analyses also suggest that effect sizes are strongly modulated by stimulus visibility: the threat-related bias was significantly larger if the SOA between stimuli and masks was >30 ms than if it was < 30ms. Critically, this suggests that unintended stimulus visibility may increase threat-related biases: many observers achieve above-chance detection of 33ms targets, as revealed by stringent signal detection measures of awareness (Pessoa, Japee, Sturman, & Ungerleider, 2006; Szczepanowski & Pessoa, 2007). In addition, we found that this effect of SOA on threat bias was greater within studies that did not implement an awareness check to verify that masking successfully eliminated stimulus visibility. Interestingly, this suggests threat related biases in the MVP paradigm could be modulated by, or perhaps even driven by residual awareness of the masked stimuli.

In MVP studies that have measured observers' awareness of masked stimuli, this is usually implemented via an independent block of trials wherein observers complete an alternative forced choice (AFC) task, such as discriminating between different masked stimuli (Carlson, Reinke, & Habib, 2009; Fox, 2002; Mogg, Bradley, & Hallowell, 1994). In general, if observers' performance does not significantly exceed chance performance in this control task, it is concluded that any threat biases obtained during the experimental trials can be attributed to processes that occur independently of awareness of the threat stimuli.

Establishing null sensitivity to stimuli via a forced choice task in this way is associated with formidable practical and conceptual issues (Wiens, 2008). For instance, awareness checks in the MVP paradigm have typically lacked statistical power, i.e. the likelihood of type II errors (failure to detect an observer's residual discrimination of target stimuli) may have been problematically high. Our meta analysis revealed that, on average, across MVP experiments, observers were classed as unaware of stimuli if 2AFC performance was less than 68%. Importantly, this permits deviations from chance performance that are moderate in magnitude (Cohen's $h = 0.38$, see Cohen, 1977), which invalidates strong statements about truly 'unconscious' processing of the masked stimuli. Another statistical issue is that if observers are selected post-hoc on the basis of chance-level performance in an awareness check, then this can bias evidence in favour of unconscious processing - reflecting a statistical principle referred to as 'regression to the mean' (Shanks, 2016). Therefore, it is important to assess not only this subset of observers, but also to consider whether individual-level awareness of stimuli predicts attentional bias across the full sample of participants. Moreover, it is important to note that only one study employed a signal detection measure (*d'*- d prime) that corrected for individual response bias (Koster, Verschuere, Burssens, Custers, & Crombez, 2007). Taken together, these limitations suggest that more rigorous methods are needed to assess awareness[1].

Another interesting question, receiving increased attention, is whether any behavioural effects of 'unconsciously' presented stimuli depend on the method used to manipulate awareness. For instance, it is possible that threat stimuli can modulate attention independently of awareness, but that the brevity of masked presentations degrades processing of the target stimuli such that any attentional modulation is reduced and hard to detect. Masking necessitates presentation times that are substantially briefer (< 40 ms) than those chosen to optimise attentional cueing effects in standard, supraliminal versions of the visual probe task (usually around 500 ms; Bar-Haim, et al., 2007). Since the presentation of stimuli in the masked version of the visual probe paradigm is an order of magnitude briefer than in the standard version, this confounds any comparison between aware and unaware processing. A more direct comparison would require that subliminal stimuli are not so temporally disadvantaged, relative to a supraliminal counterpart. Continuous flash suppression (CFS), which is an increasingly popular method in the study of unconscious processing (Sklar et al., 2012) may provide one solution to this problem. In CFS, stimuli presented to one eye can be suppressed from awareness by presenting a dynamic noise pattern to the other eye (Tsuchiya & Koch, 2005). With appropriate presentation parameters, CFS can render stimuli invisible for several seconds, allowing time for unconscious processes to engage with the suppressed stimuli. The use of CFS in a visual probe paradigm may therefore provide a more suitable comparison between unaware and aware states.

Finally, a critical conceptual issue concerns the stimulus attributes that drive the prioritisation of threat stimuli across all paradigms: standard visual probe, masked visual probe and CFS. Although it has been demonstrated that certain classes of threat stimuli, such as fearful faces, are reliably prioritized, one idea gaining traction is that this prioritization may be better explained by their low-level properties than by threat-sensitive processes (Gray, Adams, Hedger, Newton, & Garner, 2013; Hedger, Adams, & Garner, 2015b; Stein & Sterzer, 2012). For instance, Lee and colleagues (2013) found that the increased luminance contrast resulting from the greater exposure of the scleral field (eye whites) in fearful faces (relative to neutral faces) was a good predictor of enhanced performance in an attentional cuing task. More recently, work from our own lab revealed that the relationship between a face's amplitude spectrum and the human contrast sensitivity function was a better predictor of the face's detectability in masking and CFS tasks than its perceived valence or arousal (Hedger et al., 2015b). For example, fear faces have greater luminance contrast at the spatial scales humans are sensitive to than angry faces, and this predicts their higher levels of detection. This sensory advantage of the fear expression is particularly important, since fearful faces give rise to the largest, most reliable threat-related biases in the MVP paradigm of all stimulus types (Hedger et al., 2016). As highlighted in our meta analyses, it is critical that researchers provide adequate stimulus controls such that threat-related biases driven by the semantic content of stimuli (i.e. their affective content) are distinguished from effects driven by simple low-level differences between stimuli. If processing advantages are driven by low-level stimulus properties, this negates the need to invoke unconscious processes sensitive to threat.

The current study aims to address contentious or unresolved issues within the current literature. Specifically, we ask: (i) Do emotionally salient stimuli modulate attention in standard viewing conditions (i.e. with awareness of the stimuli)? (ii) Do emotional stimuli modulate attention under conditions of unawareness, as defined by stringent signal detection criteria? (iii) Are these effects modulated by the method used to render stimuli perceptually invisible? iv) Are attentional biases better explained by affective, or low-level variability across stimuli?

Method

Participants

Before recruiting participants, ethical approval for the study was obtained via the University of Southampton Research Ethics Committee (Submission ID: 17166). From our previous meta analyses (Hedger et al., 2016), we determined that 41 participants would be required to attain 95% power to detect the attentional effects observed when fear and neutral faces compete in the MVP paradigm (d_z = 0.58). For this reason, data collection was terminated when 41 undergraduate students (9 male, *M* age = 20.2 years) had completed the experiment. All observers had normal or corrected-to-normal vision.

Stimuli

Stimuli were four male facial models, taken from the NimStim face set (Tottenham et al., 2009), depicting neutral, fearful and happy expressions. All stimuli were placed within an opaque elliptical mask to eliminate external features and were equated in luminance and root mean squared (RMS) contrast. Face stimuli were presented in two configurations. *Normal* faces were presented upright with veridical contrast polarity. *Upside-down negative* faces were rotated 180 degrees with reversed contrast polarity, producing an image similar to a photographic negative (see Figure 1). These manipulations severely disrupt the recognition

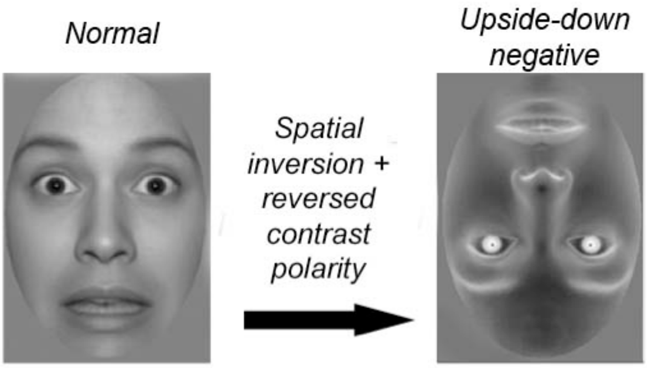

Figure 1. Example face stimulus presented in the normal and upside-down negative configuration.

and affective evaluation of facial expressions (Gray et al., 2013; Hedger et al., 2015b). Critically, however, these manipulations do not affect the low-level stimulus properties of the image: i.e., its RMS contrast, mean luminance and amplitude spectra (and therefore the energy / strength of image contours). Thus, if the valence of face images is critical in directing spatial attention, we would expect any effect of expression to be reduced or eliminated for the upside-down negative images, relative to the normal images (i.e. an interaction between expression and stimulus configuration). Conversely, if low-level properties of the stimuli explain the effect of expression, we would anticipate a similar main effect of expression for both normal and upside-down negative stimuli (i.e. no interaction between expression and stimulus configuration). All stimuli subtended 6.2 x 4.1 degrees of visual angle (DVA) at the viewing distance of 70 cm on a 1280 x 1024 pixel resolution, gamma corrected monitor. In all trials, observers viewed the display via a mirror stereoscope, and each eye's image was framed by a random dot surround (9.5 x 11.4 DVA) to control vergence.

Questionnaire Measures

Previous work suggests that attentional biases towards emotional stimuli are modulated by anxiety and related trait characteristics (Fox, 2002; Mogg & Bradley, 1999). Before the visual probe experiment, all observers completed the following measures of general and social anxiety: Trait Anxiety Inventory (STAI-T, Speilberger et al., 1983), Social Interaction Anxiety Scale (SIAS, Heimberg, Mueller, Holt, Hope, & Liebowitz, 1992) and Social Phobia Scale (SPS, Heimberg et al., 1992).

Procedure

Each trial began with the presentation of a central fixation cross whose duration was randomly sampled from the range 300 to 1000 ms to avoid anticipatory responses. Observers completed 560 trials in total. On *'signal'* trials (336 trials), pairs of face stimuli were presented to observers. On *'noise'* trials (224 trials), no face stimuli were presented to observers; intermingling signal and noise trials enabled concurrent evaluation of stimulus awareness (see 'noise trials' section). There were three presentation conditions (Figure 2).

Presentation Conditions

Standard presentation.

In the *standard* presentation condition (Figure 2a - 112 trials), two faces were presented monocularly (eye of presentation counterbalanced across trials) on either side of the fixation cross for 500 ms, whilst only the fixation cross and surround were presented to the other eye. Monocular presentation of face stimuli allowed a straightforward comparison with the CFS presentation condition. Immediately after the face presentation, a dot appeared at the location preceded by the left or right face and observers were required to report its location as quickly and accurately as possible (via left and right key button press).

Masked presentation.

In the *masked* presentation condition (Figure 2b - 112 trials), our trial sequence mirrored that of previous literature (Fox, 2002). Two face stimuli appeared binocularly either side of fixation for 17ms before being immediately replaced by two masks (patterns of high contrast ellipses) for 17ms. A 17ms SOA between face and mask has been commonly employed in previous MVP studies (Beaver, Mogg, & Bradley, 2005; Fox, 2002; Koster, Verschuere, Burssens, Custers, & Crombez, 2007; Mogg & Bradley, 1999, 2002), due to the refresh rate of standard cathode ray tube (CRT) monitors. Immediately after presentation of the mask, a dot appeared at the location preceded by the left or right face and observers were required to report its location as quickly and accurately as possible.

CFS presentation.

In the *CFS* presentation condition (Figure 2c - 112 trials), two faces were presented monocularly (counterbalanced across eyes) on either side of the fixation cross for 500ms, whilst dynamic masking patterns (refresh rate of 10Hz) were presented to the other eye, on either side of fixation. Immediately after, a dot appeared at the location preceded by the left or right face and observers were required to report its location as quickly and accurately as possible.

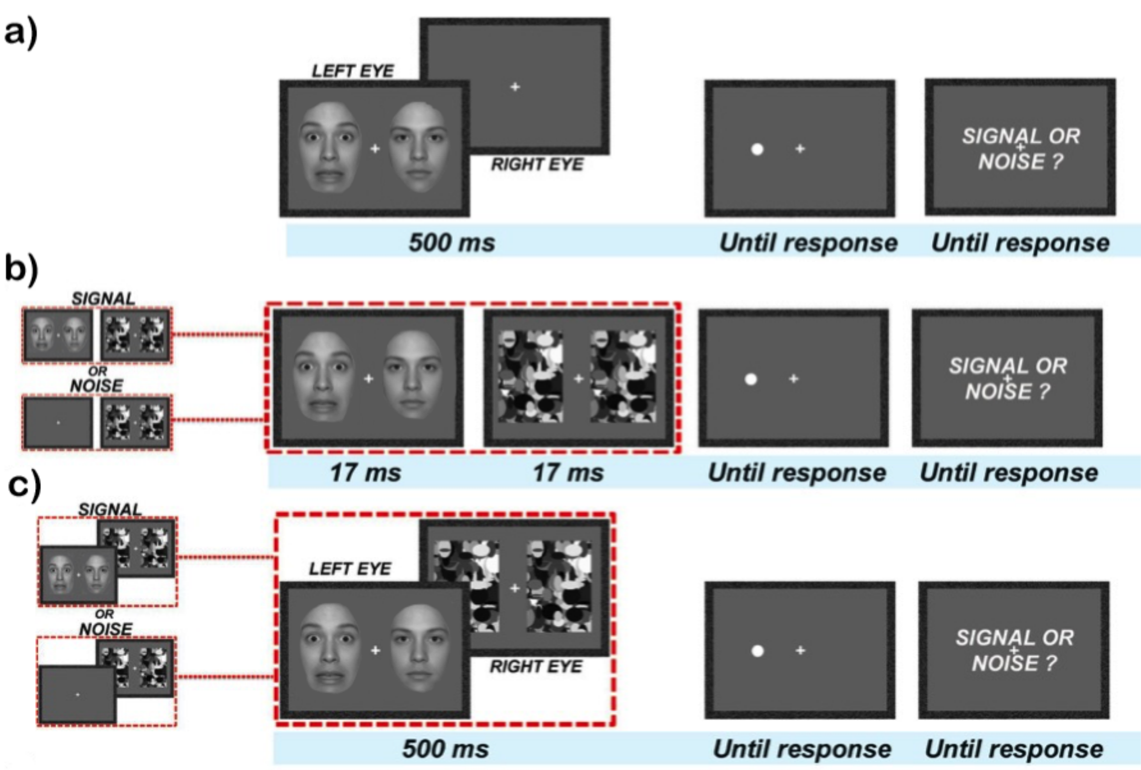

Figure 2. Schematic of trial sequences for the three presentation conditions. a) standard presentation b) masked presentation c) CFS presentation. Masked and CFS trials had an equal number of signal trials (trials where face stimuli were presented) and noise trials (trials where no face stimuli were presented) – these are shown in the leftmost panels.

Stimulus Pairing Conditions

Within each presentation condition (*standard, masked, CFS*) there were two stimulus-pairing conditions, corresponding to emotion bias trials, and face bias trials (see Figure 3).

Emotion bias

Mirroring conventional visual probe studies, *emotion bias trials* (64 trials), were designed to measure whether an emotion bias exists, i.e. a tendency to allocate attention to emotional stimuli when a neutral and an emotional stimulus compete for resources (Figure 3a). The face presentation consisted of an emotional face (32 fear, 32 happy) presented to one side of fixation and a neutral face presented to the other. Within each emotion bias pair, half of the trials were valid (subsequent probe appeared in the location of the emotional face) and half were invalid (probe appeared in the location of neutral face). These trials were repeated with face stimuli presented in both normal (16 trials, figure 3a) and upside-down negative configurations (16 trials, figure 3b).

Face bias

Face bias trials (48 trials) were designed to measure any bias for more face-like stimuli in the allocation of selective attention when normal and upside-down negative face stimuli (with matching emotional expression) compete for resources (Figure 3c). In *face bias trials*, a normal face (16 neutral, 16 fearful, 16 happy) was presented on one side of fixation and a face with the same emotional expression, but in an upside-down negative configuration was presented on the other. Within each face bias paring, half of the trials were valid (subsequent probe appeared in the location of the normal face) and half were invalid (probe appeared in the location of upside-down negative face).

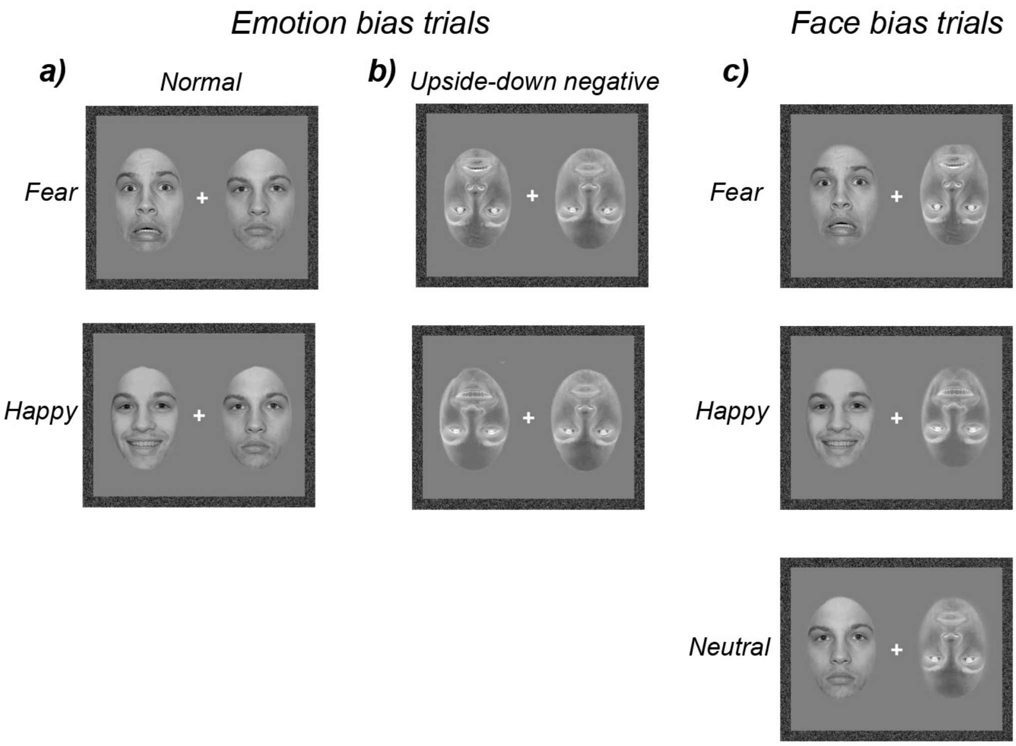

Figure 3. Schematic examples of each stimulus pairing condition. *a)* Emotion bias trials (normal configuration). A normal emotional (fear or happy) and normal neutral face were presented either side of fixation. *b)* Emotion-bias trials (upside-down negative configuration). An upside-down negative emotional (fear or happy) and upside-down negative neutral face were presented either side of fixation. c*)* Face bias trials: a normal (fear, happy or neutral) and upside-down negative face (same expression) stimulus were presented either side of fixation.

Noise Trials

50% of the trials within the CFS and masked presentation conditions (112 masked, 112 CFS) were *'noise'* trials. These trials were identical to signal trials, except no face stimuli were presented prior to the mask (for masked presentations) or to the opposite eye to the mask (for CFS presentations). If observers are unaware of the stimuli, they should perform at chance in discriminating signal trials from noise trials (Wiens, 2008). Thus, on each trial, after the observer reported the location of the probe, they were prompted to indicate whether the preceding presentation had been a ‘noise’ trial or a ‘signal’ trial (by pressing the up or down arrow key). It was clearly explained to the participants that, within those trials that contained a mask, faces were presented on only 50%, and that they had to discriminate these cases from those in which no faces were presented. Participants were also informed that there were no time constraints for this response and that they should prioritise accuracy over speed. The 224 trials for each presentation condition meant that this forced choice task had adequate (80%) power to detect even very small deviations from chance performance (Cohens *h* of 0.16 or larger).

Summary

All 41 observers completed 336 signal trials – 112 trials for each of the 3 presentation conditions (standard, masked, CFS), each comprising (i) 64 emotion bias trials: 2 emotions (fear vs. neutral, happy vs. neutral) x 2 face configurations (normal, upside-down negative) x 16 repetitions, and (ii) 48 face bias trials: 3 emotions (neutral, fear, happy) x 16 repetitions). Participants also completed 224 noise trials (112 masked, 112 CFS). The side of the emotional /upside-down negative face, the eye of face presentation, the location of the probe and the validity of the probe were counterbalanced. Trial order was randomized for each participant.

Results

Observers’ Awareness of the Face Stimuli

Following standard practice, *d'* values were computed from the difference between the z-transformed hit rates (proportion of signal trials that were correctly identified) and false alarm rates (proportion of noise trials that were incorrectly classified as signal trials). For masked presentations, *d'* was consistent with poor discrimination between signal and noise trials - at the group level, performance was not significantly better than chance ($M = 0.04$, $t(40) = 1.54$, $p = .130$). No individual observer significantly exceeded chance performance in correctly discriminating signal and noise trials (assessed via binomial test, upper binomial limit = 127 correct responses). For CFS presentations, performance was slightly higher and significantly different from zero at the group level ($M = 0.06$, $t(40) = 2.55$, $p = .015$). At the individual level, two observers performed significantly above chance in distinguishing signal and noise trials. These two observers were excluded from further analyses (with the exception of the correlation analyses shown in Figure 6). After removal of these observers, the group *d'* was not significantly different from zero for either masked ($M = 0.04$, $t(38) = 1.44$, $p = .158$) or CFS ($M = 0.04$, $t(38) = 2.01$, $p = .051$) presentations.

Visual Probe Data

Data reduction and global measures.

Preliminary inspection of the data revealed that one observer only achieved 52% accuracy in the probe discrimination task. Given the trivial difficulty of this task (discriminating left probes from right probes) we reasoned that this observer did not engage with the task requirements and thus their data were not analysed further. The remaining observers achieved near- ceiling accuracy ($M = 98.37\%$, $SD = 1.97\%$. Response times (RTs) corresponding to incorrect responses were removed (0.75% of RT data) and a log transform was applied to correct for skew. The mean log RT was calculated for each observer for each presentation condition and cue validity. Values that were more than 3 standard deviations from these means were defined as outliers and removed (1.53%). The analyses reported below were conducted on the remaining 97.72% of the RT data. After the removal of outliers, RTs were within the normal range for visual probe studies (*M* =387 ms, *SD* = 101 ms).

Emotion bias.

To test whether observers' attention was drawn to emotional faces, we calculated an emotion bias score from the *emotion bias* trials (invalid RT - valid RT) for each stimulus condition (fear or happy, within normal or upside-down negative configurations) such that positive values indicate that attention is drawn to the location of the emotional (rather than neutral) expression. These are summarised in Figure 4a. In addition, we calculated the overall emotion bias in each of the three presentation conditions (pooled across expression and configuration), these are shown in Figure 4b.

An overall emotion bias was detected for standard presentations (*M* = 19.79 ms, *t* (37)=2.33, *p*=.025), corresponding to a modest effect size (d_z = 0.38, 95% CI [0.05 0.71]). However, in the masked and CFS conditions, no overall emotion bias was detected (masking: *M*= -7.26 ms, *t* (37) =-0.87, *p*=.382, CFS: *M*= 0.07 ms, *t* (37) =0.01 *p*=.991), and the effect sizes were small (masking: d_z = -0.14, 95% CI [-0.46 0.17], CFS: d_z = 0.00, 95% CI [-0.32 0.32]). When comparing the magnitude of the emotion bias across conditions, a main effect of presentation condition was detected (*F* (2,74) = 3.096, *p* =.046). Thus, our data suggest that observers' attention was drawn towards emotional stimuli under standard presentation, in which stimuli were fully visible, but not in masking or CFS trials.

We can consider whether the emotion bias was modulated by expression (fear vs. happy) or configuration (normal vs. upside-down negative); if attentional allocation was driven by affective content, (i.e. the meaning of the stimuli) then we expect a larger emotion bias in the normal than the upside-down negative configuration, because expressions are harder to discriminate in the upside-down negative configuration. There was no significant interaction between expression and stimulus configuration in any presentation condition (2-way ANOVAs, all *p*-values > .45). Importantly, this suggests that facial expression had no effect on attentional allocation beyond that explained by basic low-level variability between expressions. In fact, the emotion bias in the standard presentation condition (widely reported in previous literature: Bar-Haim et al., 2007) was *smaller* for normal than upside-down negative stimuli (normal: $M = 18.56$, upside-down negative: $M = 21.02$, $t(37) = -0.17$, $p = .863$).

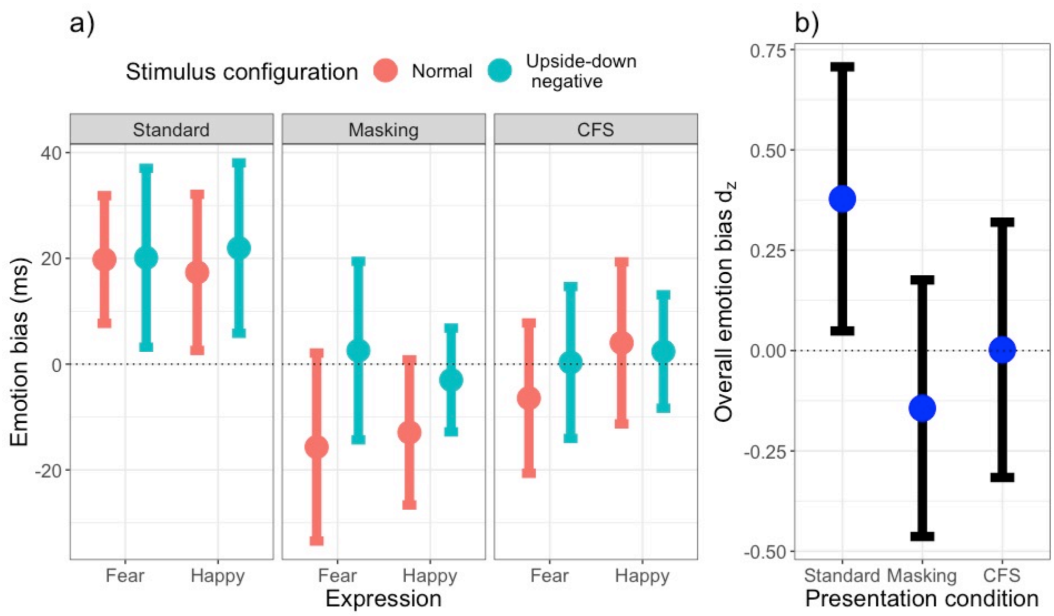

Figure 4. Attentional biases in emotion bias trials. a) Emotion bias (invalid RT - valid RT) plotted as a function of expression, stimulus configuration (normal, upside-down negative) and presentation condition. Error bars are +/- 1 SE. b) The overall emotion bias, expressed as Cohen's d_z is plotted as a function of presentation condition. Error bars are 95% confidence intervals.

Face bias.

To determine whether observers' attention was directed to more face-like stimuli, i.e. normal faces, as opposed to upside-down negative faces, we calculated a face bias score from *face bias* trials (invalid RT - valid RT) for each stimulus condition, such that positive values indicate that attention is drawn to the location of the normal face. The resultant face biases are summarised in figure 5a. Figure 5b shows the overall face bias effect size in each presentation condition (pooled across all stimuli). We detected a significant face bias in standard trials (M = 18.33 ms, t (37) = 2.19, dz = 0.36, [0.03 0.68], p = .035) and in CFS trials (M = 28.39 ms, t (37) = 2.98, dz = 0.49, [0.15 0.82], p = .005) but not in masking trials (M = -11.01 ms, t (37) = -1.48, dz = - 0.24, [-0.56 0.08], p =.146). A significant effect of presentation condition (F (2,74) = 5.64 p =.005) was detected. Post hoc tests revealed that both standard (p =.015) and CFS (p =.003) presentations yielded larger face biases than masked presentations. We can ask whether these biases towards normal faces are modulated by expression. However, there was no significant main effect of expression F (2,74) = .099, p = .906, or interaction with presentation condition F (4, 148) = .92, p =.455. We detected no effect of expression within any presentation condition (one-way ANOVAs, all p-values > .377). Despite observers indicating low levels of overall sensitivity to the presence v absence of stimuli, one possibility is that small differences in sensitivity to normal vs. upside-down negative faces (e.g. Jiang, Costello & He, 2007) could account for the face biases in CFS trials. Since both normal and upside-down negative stimuli are presented in face bias trials, we tested this possibility by using the signal detection data from emotion bias trials task to predict face bias. Since 'noise' trials only differ at the level of presentation condition (CFS, masked), it is not possible to calculate d' seperately for individual trial types (e.g. normal and upside-down negative trials). For this reason, we used hit rate as our measure of sensitivity in this analysis. We detected no association between differential hit rate (normal – upside-down negative hit rate) and face-bias magnitude (r (36)=-.06, [-.38 .261] , p=.703.

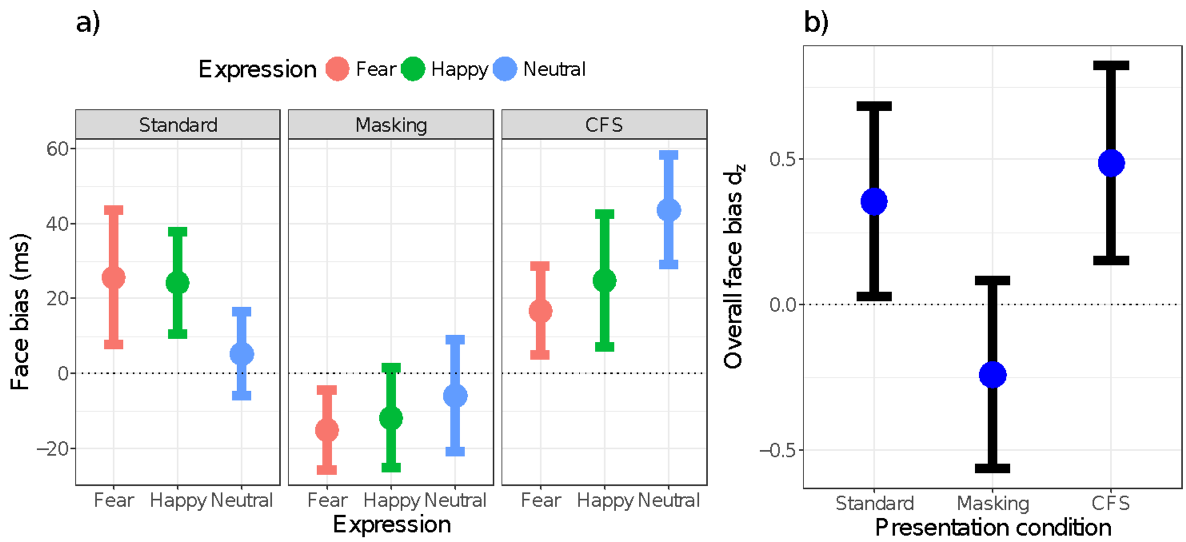

Figure 5. Attentional biases in face bias trials. a) Face bias scores (invalid RT - valid RT) plotted as a function of expression and presentation condition. Error bars are +/- 1 SE. b) The overall face bias, expressed as Cohen's d_z, as a function of presentation condition. Error bars are 95% confidence intervals.

Modelling of Visual Probe Data

To better understand the stimulus attributes that predict attentional capture, and their relative importance, we developed a simple model to explain not just the mean difference between valid and invalid RTs within each stimulus pairing (which discards some, potentially informative RT information) but the full set of raw response times. In the visual probe paradigm, each trial represents a competition for attention between two stimuli (A and B). We use the term 'salience' (S) to represent the capacity of each stimulus to capture attention, i.e. S_A and S_B. If $S_A > S_B$ then, when a probe follows stimulus A, reaction times (RT_A) will be shorter, on average, than when it follows stimulus B (RT_B). Reaction times are well approximated by a log-normal distribution, and so we model the probability distribution of reaction times as:

$$p(\log(RT_A)=t)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{(t-\mu-S_A+S_B)}{2\sigma^2}}$$

where μ and σ give the mean and standard deviation of the baseline distribution of log (RT), i.e. with two equally salient stimuli.

We considered different stimulus attributes that might modulate stimulus salience. These included *emotional* content (emotional vs. neutral), *threat* (the additional salience of fearful faces, relative to happy faces) and *configuration*, i.e. similarity to a normal face (normal vs. upside-down negative). We tested models that included only additive (e.g. *emotion* + *configuration*) or also interactive (e.g. *emotion* * *configuration*) combinations of these variables. These parameters were used to define the relative salience of stimuli, where a neural upright face is given a nominal saliency of 0. For example, if the salience of a stimulus were independently predicted by emotion and configuration (Model 5, Table 1), then the (relative) salience of normal fearful and happy faces is given by the *emotion* coefficient, and the salience of neutral upside-down negative faces is given by the *configuration* coefficient. A positive *emotion* coefficient indicates that emotional (fear and happy) faces are more salient than neutral faces and a negative *configuration* coefficient indicates that normal faces are more salient than upside-down negative faces. Interaction terms allow us to model the situation in which the effect of emotion or expression differs between normal and upside-down negative configurations.

Table 1.

Summary of Tested Models and their Parameters

Model	Parameters
1	*(Null model)*
2	*configuration*
3	*emotion*
4	*emotion, threat*
5	*configuration, emotion*
6	*configuration, emotion, threat*
7	*configuration, emotion, configuration*emotion*
8	*configuration, emotion, threat, config*emotion, config*threat.*

Because our models differed in complexity (the number of free parameters), leave one out cross-validation (LOO xval) was used to evaluate the generalisation performance of all models and avoid over-fitting (see Supplementary Material S1). In this method, a model is fit to *N*-1 observers (training data) and the fitted values are used to predict the data from the 'left out' observer (test data). The performance of the model in predicting the new data (in terms of error) directly reflects the generalisation performance of the model in predicting new 'unseen' data. One appealing property of LOO xval is that, unlike model performance indices such as Bayesian Information Criterion (BIC) or Akaike's information Criterion (AIC), there is no need to apply (ad-hoc) criteria to determine whether a more complex model's improved fit (in terms of likelihood, or other goodness-of-fit metric) is justified by the increased number of free parameters. Instead, the LOO method will naturally reveal the number of parameters required to model the signal (but not the noise) within the dataset: Unnecessarily complex models are implicitly penalised by this procedure, since they 'overfit' to the training data and therefore will have lower performance in predicting the left out (test) data. The results of the LOO xval procedure are displayed in Supplementary Material S2 (this figure illustrates that several models perform worse at cross-validation than the null model, demonstrating the penalty associated with over-fitting). The fitted parameters for all models are summarised in Supplementary Material S3.

The model that best explains stimulus salience, and thus participants' RTs, varied according to presentation condition (see Table 2). Under standard presentations, where observers were fully aware of the face stimuli, the data were best explained by a model of salience that included both *emotion* and *threat* (model 4). In other words, emotional stimuli attracted attention more effectively than neutral stimuli, and this effect was increased for threat-relevant fear faces. Importantly, models involving interactions between *configuration* and *emotion* or *threat* did not improve on this model. In other words, under conditions of full awareness, there is no evidence that emotional stimuli have increased salience, beyond that determined by their low-level image properties. The modelling results are broadly in line with the traditional visual-probe analyses reported above, although they additionally have the increased sensitivity to reveal the increased salience of fear, relative to happy faces.

Table 2.

Best Fitting Models for Each Presentation Condition

Presentation Condition	Best Model	Fitted coefficients *M (SD)*
Standard	4	*emotion*: 5.98 (0.56)
		threat: 2.14 (0.64)
Masking	2	*configuration*: 5.31 (0.42)
CFS	2	*configuration*: -6.02 (0.37)

For both masked and CFS presentations, the best model of participants' RTs included only the stimulus configuration (model 2). More complex models, involving stimulus emotion or threat did not have greater cross validation performance than the null model. This, alongside the previous analyses, suggests that participants' responses are not affected by facial emotion when faces are presented outside of awareness via backward masking or CFS.

Association With Awareness and Anxiety Measures

Our previous work (Hedger et al., 2016) suggests that some emotion-related biases found in backward masking and CFS paradigms may be due to, or modulated by observers' awareness of the stimuli. To determine whether such an effect exists in the current data, we examined the relationship between attentional bias and awareness of the stimuli (as indexed by performance in the 'signal' vs. 'noise' discrimination task) in the masked and CFS presentation conditions. For each observer, we computed a single attentional bias score, collapsed across all stimulus types, and a single *d'* score for stimulus awareness, collapsed across CFS and masked presentations.

Performance in the awareness task (i.e. the ability to distinguish 'signal' from 'noise' trials) was significantly and positively correlated with attentional bias (F (1,38) = 4.693, R^2 = .086, p=.037), as shown in Figure 6a, suggesting that attentional biases are inflated when observers have some awareness of the stimuli. Notably, the best-fit line passes very close to (0,0), suggesting that awareness of the stimuli not only increases attentional bias, but may be required for attentional bias effects to occur. When the data were split by stimulus pairing, this revealed that the association between awareness and attentional bias was mostly driven by the emotion bias $F(1,38)$ = 3.311, R^2 = .080, p =.077 (Figure 6b) and that the face bias had a weaker association with awareness $F(1,38)$ = 2.003, R^2 = .025, p =.166 (Figure 6c).

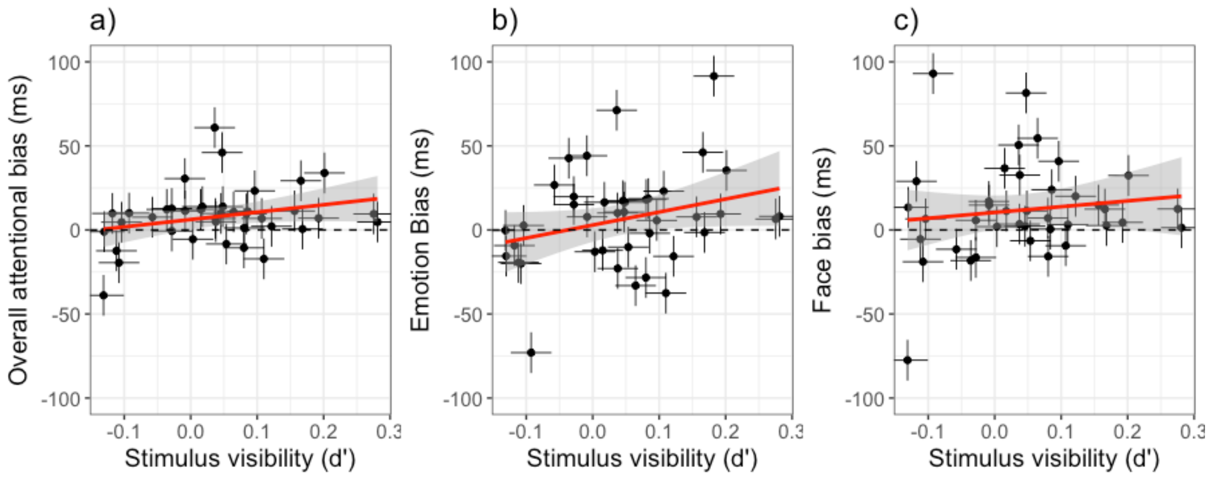

Figure 6. a) Association between overall *d'* and overall attentional bias score. b) Association between *d'* and emotion bias score. c) Association between *d'* and face bias score. Red lines are the least squares fit to the data, shaded region is +/- 1 SE.

As shown in supplementary material S4, our measures of anxiety (STA-T, SPI, SIAS) were highly correlated (all Pearson's' $r > .43$ all ps <.01), indicating good reliability of these measurements. Our primary research questions were unrelated to individual differences (e.g. participants' levels of anxiety). However, previous research has suggested that biases toward emotional stimuli are inflated amongst anxious observers (Bar- Haim et al, 2006). For this reason, we performed correlation analyses to examine the association between attentional bias and anxiety measures (STAI-T, SPS, SIAS) in each presentation condition. No correlations were detected in any presentation condition for either emotion or face bias scores (all *p*-values >.05). These correlations and their associated confidence intervals can be found in supplementary material S5.

Discussion

Our experimental design allowed us to assess attentional orienting to neutral and emotional face stimuli under different conditions of awareness. A few key findings emerged: (i) In the standard, supraliminal paradigm, we found evidence that emotional faces attract attention when competing with neutral faces. Our stimulus configuration manipulation (normal v upside-down negative) allowed us to determine the extent to which this is driven by the low-level image properties of the stimuli vs. their recognisable emotional content. In fact, the effect of emotion was slightly larger for upside-down negative faces, suggesting that attentional allocation within our visual probe task was not driven by recognisable emotion. No emotion biases were found when stimuli were presented outside of awareness via masking or CFS. (ii) Normal faces attracted attention over upside-down negative faces within the standard and CFS conditions, suggesting a preference for more natural, face-like stimuli. (iii) Attentional effects were predicted by observers' awareness of the stimuli, suggesting that attentional biases are modulated, or even driven by awareness.

Attentional Capture by Emotionally Salient Stimuli

For standard, 500ms supraliminal trials, we observed attentional biases towards emotionally salient stimuli. Importantly, these effects were not reduced when the stimuli were presented in the upside-down negative condition (in fact, they increased slightly). This suggests that the apparent effect of emotion on attentional allocation within the standard (conscious) visual probe paradigm was driven by low-level stimulus factors (e.g. luminance contrast), rather than emotional valence per se; the upside-down negative faces have vastly reduced recognisable emotional content (Gray et al., 2013). This finding suggests that previous reports of an attentional bias towards emotional stimuli may be attributable, at least in part, to low-level differences between stimuli. Fearful and happy facial expressions in particular tend to have more contrast energy at the spatial scales humans are sensitive to, relative to neutral faces (Hedger et al., 2015). Thus, a processing advantage for these expressions is predicted on the basis of simple sensory factors alone.

When stimuli were suppressed from awareness via masking and CFS, emotion bias effects were small and insignificant. However, as with any null result, it is worth discussing possible sources of a type 2 error. On a statistical level, it is important to note that power was high and the sample size calculations were based on a large body of previous literature (see 'participants' section). Secondly, in relation to the sample characteristics, the mean trait anxiety level was relatively high (*M*= 41.46) and above the value expected to produce detectable biases towards threatening stimuli under subliminal presentations (Hedger et al., 2016). Unlike studies that have solely used masking to manipulate awareness, it is unlikely that null effects in unconscious presentations can be explained by simple restrictions on presentation time, since this was equated in normal and CFS trials (500 ms). Importantly, we detected significant attentional bias effects under standard presentation conditions, suggesting that the task, in itself, was a sensitive measure of attentional allocation. Our results are consistent with studies that have failed to detect evidence of emotional modulation of attention under masking (Fox, Cahill, & Zougkou, 2010; Koster et al., 2007) and CFS (Hedger, Adams, & Garner, 2015a).

The Thorny Issue of Low-level Confounds

In our study, we observed equivalent attentional cuing effects for normal and upside-down negative stimuli and thus concluded that variability between stimuli in low-level image properties drives attentional cuing. It is worth discussing some competing explanations for our findings.

First, we consider whether normal and upside-down negative stimuli share some property (other than luminance contrast) that drives the attentional effect. For instance, upside-down negative stimuli could retain some (reduced) recognisable emotional content, and this attracts attention to the same extent as the emotion as content within normal stimuli. We think this is unlikely for two reasons. Firstly, upside-down negative stimuli actually give rise to slightly *larger* attentional effects than normal stimuli (rather than smaller effects, see Figure 4). Secondly, we have previously demonstrated that upside-down negative facial expressions are not evaluated as being significantly different from neutral valence according to an Extrinsic Affective Simon Task (EAST). Specifically, the disruption caused by this manipulation was found to not simply reduce the magnitude of valence judgements, but generated a qualitatively pattern of effects (see Gray et al., 2013 experiment 2).

A second possibility is that the similar effects observed for normal and upside-down negative stimuli are not driven by a single, shared mechanism, but by two different mechanisms that happen to produce effects of similar magnitude. For example, might it be that normal faces capture attention via their emotional content, whereas upside-down negative faces capture attention via their low-level image properties? Again, we believe that this explanation is unlikely. Firstly, it contradicts standard experimental logic – if one manipulates a potentially important experimental variable (e.g. emotional content), and finds no effect on the dependent variable (e.g. attentional allocation), the standard conclusion is that the variable is not as important as previously hypothesised. Second, we have previously demonstrated that stimulus detection (across both normal and upside-down negative faces) is much better predicted by variations in luminance contrast than by variations in perceived valence or arousal (Hedger et al 2015b experiment 2). Moreover, a quantitative analysis of previous literature also reveals that the affective content of a stimulus is a poor predictor of its ability to capture attention in the MVP paradigm: our meta-analysis revealed that fearful faces are the only class of threat stimuli that reliably generate threat-related biases in the MVP paradigm. Angry faces, in contrast, yield small, non-significant effects (Hedger et al., 2016). Angry faces signal a direct threat to the observer, whereas fearful faces indicate the presence of a proximal threat. If threat directs attention within VP paradigms, this should be apparent for indicators of proximal threat (i.e. fearful faces), and direct threat (i.e. angry faces). Instead, parsimony favours an account based on the sensory advantages of the fear expression over the neutral or angry expression (Hedger et al., 2015b, Lee et al., 2013). Another potential objection could be that the unfamiliar/unusual quality of upside-down negative faces might drive attention. Crucially, it is worth noting that this concern would not apply to emotion bias trials, since both emotional and neutral face stimuli were presented in the upside-down negative configuration on these trials. We also explicitly test for the possibility that 'unfamiliar' upside-down negative stimuli attract attention over the more 'familiar' normal stimuli in face-bias trials. However, the opposite effect was observed: attention was instead drawn to faces in the normal configuration. Moreover, rendering our upside-down negative stimuli 'unfamiliar' is a necessary and desired effect of the manipulation - if the face stimuli were recognisable in the upside-down negative condition, they would not provide a valid control for variations in low-level image properties.

Other objections to the conclusions drawn from our upside-down negative manipulation are more philosophical. Some authors have proposed that low-level image properties may drive efficient detection precisely because of their emotionality (Frischen, Eastwood, & Smilek, 2008). Under this line of reasoning, labelling the low-level variability between stimuli as a 'confound' is problematic, because the communicated emotion is defined by its low-level properties (e.g. patches of high contrast signal fear). This would undermine the idea of attempting to control for low-level stimulus properties in any perceptual experiment. Like others (e.g. Becker, Anderson, Mortensen, Neufeld, & Neel, 2011), we believe that this position is unfalsifiable and unsound. Clearly, not all high-contrast stimuli are fear-inducing, and as discussed above, not all threatening stimuli are high-contrast. Further, if one group of researchers holds the view that fear faces attract attention because of their threat relevance and another claims that fearful faces attract attention due to low-level image properties, then the debate can only move forward by designing experiments that attempt to resolve between these competing possibilities.

Considering all evidence, we believe that simple low-level variability between stimuli provides the most parsimonious account of the attentional effects that we observe. Other explanations are either i) harder to support with the available data, or ii) require additional assumptions.

Implications for Clinical Work.

Failing to control for, or characterize low-level stimulus properties can have serious implications. Consider populations who might be expected to show *diminished* threat processing, such as patients with a recent brain injury (Tsuchiya et al., 2009), or individuals who have received an intervention to alleviate anxiety symptoms (Murphy, Downham, Cowen, & Harmer, 2008). An apparent threat-related bias in these populations may be wrongly as interpreted as being indicative of 'unimpaired threat processing' or a 'failed intervention'. In reality, it may be that these observers simply have normal sensitivity to the low-level variability between neutral and emotional stimuli. Indeed, much of the unaccounted-for variation in the efficacy in a behavioral intervention for anxiety such as attentional bias modification (Mogg & Bradley, 2018) could be explained by low-level variability between stimulus categories.

Attentional Preference for 'Face Like' Stimuli

We observed evidence for attentional biases toward normal faces when competing with manipulated, upside-down negative faces in both standard and CFS trials. The latter finding is consistent with a large body of work from the breaking continuous flash suppression (bCFS) literature, which has consistently demonstrated that upright faces break CFS suppression faster / more frequently than inverted faces (Jiang, Costello, & He, 2007;

Stein & Sterzer, 2012). In addition, there is evidence from fMRI that face-selective regions of the temporal cortex are differentially activated by upright vs. inverted faces even when these are presented under CFS (Jiang & He, 2006). The present study extends this literature by providing the first behavioural evidence that face-like stimuli attract spatial attention under CFS. Based on the present data, inferring a causal link between preferential detection of upright, normal faces (in bCFS studies) and attention to normal faces (in the present study) would be premature. Future work could test for the presence and directionality of such a causal relationship.

Why is it that CFS seems to spare selective attention to face like configurations (in face bias trials), but not emotional expressions (in emotion bias trials)? One possibility is that discriminating a face from a non-face (a coarse, basic-level classification) is easier than discriminating different expressions (a finer, sub-ordinate classification) and may thus be less affected by degradation associated with CFS suppression. This sensitivity to upright faces cannot be explained by low-level stimulus properties such as contrast and spatial frequency profile, which are preserved after spatial and contrast inversion. Instead, the prioritised detection of upright faces appears to reflect some higher-level 'face-sensitive' process. An alternative explanation is that the preferential processing of upright faces does not reflect face-sensitive processes, but rather the fact that 'top heavy' patterns in general are more easily detectable, since humans have a robust upper hemifield advantage in basic visual sensitivity (Skrandies, 1987). In fact, recent work has shown that upright 'protofacial' stimuli (a simple triangular configuration of dots, resembling the position of the eyes and mouth) break CFS more rapidly than their inverted counterparts (Akechi et al., 2015).

Future work should aim to dissociate effects driven by face sensitive processes from effects driven by simple differences in sensitivity in the upper and lower hemifield. Clearly, these two possibilities have drastically different implications for the level and complexity of visual processing that transpires without awareness. This, in turn, suggests caution when inferring high level processing based on a preference for normal, upright face configurations.

Implications for Paradigms used to Manipulate Awareness

A recent concern, that has been raised by many, is whether the perceptual suppression induced by techniques such as masking and CFS are functionally similar to those that may occur under natural viewing conditions (Blake, Brascamp, & Heeger, 2014; Hesselmann & Moors, 2015). If they are not, then studies employing these techniques may tell us about the peculiarities of the techniques used, rather than revealing any characteristics of unconscious processing that generalise to natural viewing conditions. A related concern is that conclusions emanating from different paradigms used to manipulate awareness may not generalise to one another, since they do not index the same level of unconscious processing (Breitmeyer, 2015; Dubois & Faivre, 2014). This entails that a null effect in one paradigm does not necessarily entail the absence of unconscious processing, since affirmative findings may be found with a different paradigm. Our findings strengthen these concerns. For instance, based on our data from the standard and masked presentations alone, one could conclude that an attentional preference for face-like configurations (in face bias trials) depends on their conscious registration. By contrast, when these data are considered alongside the CFS data, one could conclude that the absence of such effects may be due to the methodological limitations of the masking paradigm. For instance, it may be the case that face-sensitive processes simply require a more sustained and robust visual signal than is supported by very brief, masked presentation. Similarly, the absence of an emotion bias in CFS or masked presentations does not necessarily imply that emotional stimuli fail to modulate attention under all conditions of unawareness. For instance, Faivre, Berthet and Koudier (2012) found that affective priming was eliminated when primes were presented under CFS, but robust priming effects were observed when primes were rendered indiscriminable by crowding. The study of unconscious processing is thus highly susceptible to the error of ‘denying the antecedent’ when interpreting null effects.

Implications for Assessment of Awareness

If attentional cuing operates independently of awareness of the cuing stimuli, we should expect no association between discrimination of stimulus presence and the magnitude of the attentional cuing effect. Instead, our data reveal that increased stimulus awareness (as assessed by *d’*) predicted increased attentional biases, despite the limited range of *d’* values and our sample’s relatively low level of sensitivity. Recent research employing stringent signal detection measures of awareness have revealed that observers are more capable than previously assumed at detecting brief, masked signals. In fact, one study has shown that the majority of observers can reliably detect images of fearful faces that are masked after 25, or even 17 ms (Szczepanowski & Pessoa, 2007). Although these deviations from chance performance were small, they are non-trivial in the context of the attentional effects emanating from the masked visual probe paradigm, which are also very small. This, taken together with our own data, illustrates the importance of providing sensitive, well-powered and objective awareness measures.

Implications for Emotion Theory

Several dominant neurocognitive theories of emotion assume independence of affective processing and awareness. Various ‘dual pathway’ models rest on the assumption that processing of affective visual stimuli involves a separable sub-cortical visual pathway that bypasses the visual cortex and projects affective information rapidly to emotionally-responsive structures (e.g. the amygdala) independently of awareness. The first explicit model of this kind was formulated as early as 1885 (Lange, 1885) and adaptations of this idea have been presented more recently (LeDoux, 1996; Tamietto and de Gelder, 2010). Clearly, it would be rash to challenge the neuroanatomical aspects of such theories on the basis of our behavioural data. However, regardless of whether such a pathway exists, we find no evidence that it supports the preferential processing of threat stimuli in the absence of awareness. This accords with recent suggestions that the proposed subcortical pathway is highly unlikely to have the computational properties required to perform processes such as object identification, which would be required in order to differentiate threatening from nonthreatening signals (Cauchoix & Crouzet, 2013).

Conclusion

In conclusion, our data suggest that attentional capture by emotionally salient stimuli is predicted by awareness. We detected attentional cuing effects under normal viewing conditions, but not under two different conditions of unawareness. Moreover, we provide direct evidence that an observer's awareness of stimuli predicts the magnitude of attentional cuing effects. Finally, even under full awareness, we found that attentional cuing by emotionally salient stimuli was fully accounted for by low-level stimulus confounds. When considered alongside our meta-analysis, these findings could motivate a reinterpretation of previous literature and stimulate further well-controlled studies on the relationship between emotion processing, attention and awareness.

References

Akechi, H., Stein, T., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2015). Preferential awareness of protofacial stimuli in autism. *Cognition*, *143*, 129–134. http://doi.org/10.1016/j.cognition.2015.06.016

Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia*, *40*(7), 817-826.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. http://doi.org/10.1037/0033-2909.133.1.1

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: happy faces, not angry faces, are more efficiently detected in single- and multiple-target visual search tasks. *Journal of Experimental Psychology. General*, *140*(4), 637–659.

Beaver, J. D., Mogg, K., & Bradley, B. P. (2005). Emotional conditioning to masked stimuli and modulation of visuospatial attention. *Emotion (Washington, D.C.)*, *5*(1), 67–79. http://doi.org/10.1037/1528-3542.5.1.67

Blake, R., Brascamp, J., & Heeger, D. J. (2014). Can binocular rivalry reveal neural correlates of consciousness? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1641), 20130211. http://doi.org/10.1098/rstb.2013.0211

Breitmeyer, B. G. (2015). Psychophysical "blinding" methods reveal a functional hierarchy of unconscious visual processing. *Consciousness and Cognition*, *35*, 234-250. https://doi.org/10.1016/j.concog.2015.01.012

Carlson, J. M., & Reinke, K. S. (2008). Masked fearful faces modulate the orienting of covert spatial attention. *Emotion (Washington, D.C.)*, *8*(4), 522–529. http://doi.org/10.1037/a0012653

Carlson, J. M., Reinke, K. S., & Habib, R. (2009). A left amygdala mediated network for rapid orienting to masked fearful faces. *Neuropsychologia*, *47*(5), 1386–1389. http://doi.org/10.1016/j.neuropsychologia.2009.01.026

Cauchoix, M., & Crouzet, S. M. (2013). How plausible is a subcortical account of rapid visual recognition? *Frontiers in Human Neuroscience*, *7*, 39.

Chica, A. B., Martín-Arévalo, E., Botta, F., & Lupiáñez, J. (2014). The Spatial Orienting paradigm: How to design and interpret spatial attention experiments. *Neuroscience & Biobehavioral Reviews*, *40*, 35–51. http://doi.org/10.1016/j.neubiorev.2014.01.002

Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences.* New York: Routledge.

Dennett, D. C. (1993). *Consciousness Explained*. Penguin UK.

Dubois, J., & Faivre, N. (2014). Invisible, but how? The depth of unconscious processing as inferred from different suppression techniques. *Consciousness Research*, *5*, 1117. http://doi.org/10.3389/fpsyg.2014.01117

Faivre, N., Berthet, V., & Kouider, S. (2012). Nonconscious influences from emotional faces: A comparison of visual crowding, masking, and continuous flash suppression. *Frontiers in Psychology, 3,* 129. http://dx .doi.org/10.3389/fpsyg.2012.00129

Fox, E. (2002). Processing emotional facial expressions: The role of anxiety and awareness. *Cognitive, Affective & Behavioral Neuroscience*, *2*(1), 52–63.

Fox, E., Cahill, S., & Zougkou, K. (2010). Preconscious processing biases predict emotional reactivity to stress. *Biological Psychiatry*, *67*(4), 371–377. http://doi.org/10.1016/j.biopsych.2009.11.018

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological Bulletin*, *134*(5), 662–676.

Garrido, M. I. (2012). Brain Connectivity: The Feel of Blindsight. *Current Biology*, *22*(15), R599–R600. http://doi.org/10.1016/j.cub.2012.06.012

Gray, K. L. H., Adams, W. J., Hedger, N., Newton, K. E., & Garner, M. (2013). Faces and awareness: low-level, not emotional factors determine perceptual dominance. *Emotion (Washington, D.C.)*, *13*(3), 537–544. http://doi.org/10.1037/a0031403

Hedger, N., Adams, W. J., & Garner, M. (2015a). Autonomic arousal and attentional orienting to visual threat are predicted by awareness. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(3), 798–806. http://doi.org/10.1037/xhp0000051

Hedger, N., Adams, W. J., & Garner, M. (2015b). Fearful faces have a sensory advantage in the competition for awareness. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(6), 1748–1757. http://doi.org/10.1037/xhp0000127

Hedger, N., H, L., Garner, M., & Adams, W. J. (2016). Are Visual Threats Prioritized Without Awareness? A Critical Review and Meta-Analysis Involving 3 Behavioral Paradigms and 2696 Observers. *Psychological Bulletin*, No Pagination Specified. http://doi.org/10.1037/bul0000054

Heimberg, R. G., Mueller, G. P., Holt, C. S., Hope, D. A., & Liebowitz, M. R. (1992). Assessment of anxiety in social interaction and being observed by others: The social interaction anxiety scale and the Social Phobia Scale. *Behavior Therapy*, *23*(1), 53–73. http://doi.org/10.1016/S0005-7894(05)80308-9

Hesselmann, G., & Moors, P. (2015). Definitely maybe: can unconscious processes perform the same functions as conscious processes? *Consciousness Research*, 584. http://doi.org/10.3389/fpsyg.2015.00584

Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological science*, *18*(4), 349-355.

Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: dissociating subsystems for facial-information processing. *Current Biology: CB*, *16*(20), 2023–2029. http://doi.org/10.1016/j.cub.2006.08.084

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends in Cognitive Sciences*, *9*(8), 381–388. http://doi.org/10.1016/j.tics.2005.06.012

Koch, C. (2004). *The quest for consciousness: A neurobiological approach* (1st ed.). Denver, CO: Roberts & Company Publishers.

Koster, E. H. W., Verschuere, B., Burssens, B., Custers, R., & Crombez, G. (2007). Attention for emotional faces under restricted awareness revisited: do emotional faces automatically attract attention? *Emotion (Washington, D.C.)*, *7*(2), 285–295. http://doi.org/10.1037/1528-3542.7.2.285

Lange, C., 1885. *The mechanism of the emotions.* In: Dunlap, E. (Ed.), The Emotions. Williams & Wilkins, Baltimore, Maryland, pp. 33–92.

Lapate, R. C., Rokers, B., Li, T., & Davidson, R. J. (2013). Nonconscious Emotional Activation Colors First Impressions A Regulatory Role for Conscious Awareness. *Psychological Science*, 956797613503175. http://doi.org/10.1177/0956797613503175

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster.

Lee, D. H., Susskind, J. M., & Anderson, A. K. (2013). Social transmission of the sensory benefits of eye widening in fear expressions. *Psychological Science*, *24*(6), 957-965.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy*, *36*(9), 809–848. http://doi.org/10.1016/S0005-7967(98)00063-1

Mogg, K., & Bradley, B. P. (1999). Orienting of Attention to Threatening Facial Expressions Presented under Conditions of Restricted Awareness. *Cognition and Emotion*, *13*(6), 713–740. http://doi.org/10.1080/026999399379050

Mogg, K., & Bradley, B. P. (2002). Selective orienting of attention to masked threat faces in social anxiety. *Behaviour Research and Therapy*, *40*(12), 1403–1414.

Mogg, K., Bradley, B. P., & Hallowell, N. (1994). Attentional bias to threat: roles of trait anxiety, stressful events, and awareness. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *47*(4), 841–864.

Morris, J. S., Öhman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating 'unseen' fear. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(4), 1680–1685.

Murphy, S. E., Downham, C., Cowen, P. J., & Harmer, C. J. (2008). Direct effects of diazepam on emotional processing in healthy volunteers. *Psychopharmacology*, *199*(4), 503–513. http://doi.org/10.1007/s00213-008-1082-2

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466-478.

Ohman, A., & Soares, J. J. (1994). 'Unconscious anxiety': phobic responses to masked stimuli. *Journal of Abnormal Psychology*, *103*(2), 231–240.

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral Cortex (New York, N.Y.: 1991)*, *16*(3), 366–375. http://doi.org/10.1093/cercor/bhi115

Sekar, K., Findley, W. M., Poeppel, D., & Llinás, R. R. (2013). Cortical response tracking the conscious experience of threshold duration visual stimuli indicates visual perception is all or none. *Proceedings of the National Academy of Sciences*, *110*(14), 5642–5647. http://doi.org/10.1073/pnas.1302229110

Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review*, *24*(3), 752–775.

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences*, *109*(48), 19614–19619. http://doi.org/10.1073/pnas.1211645109

Skrandies W. (1987). “The upper and lower visual field of man: Electrophysiological and functional differences". In: Ottoson D. (Ed.), *Progress in Sensory Physiology*. *Berlin: Springer;*

Stein, T., & Sterzer, P. (2012). Not just another face in the crowd: detecting emotional schematic faces during continuous flash suppression. *Emotion (Washington, D.C.)*, *12*(5), 988–996. http://doi.org/10.1037/a0026944

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Szczepanowski, R., & Pessoa, L. (2007). Fear perception: Can objective and subjective awareness measures be dissociated? *Journal of Vision*, *7*(4), 10–10. http://doi.org/10.1167/7.4.10

Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*(10), 697–709. http://doi.org/10.1038/nrn2889

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*(8), 1096–1101. http://doi.org/10.1038/nn1500

Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., & Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nature Neuroscience*, *12*(10), 1224–1225. http://doi.org/10.1038/nn.2380

Vuilleumier, P., & Schwartz, S. (2001). Emotional facial expressions capture attention. *Neurology*, *56*(2), 153-158.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., … Johnstone, T. (2004). Human Amygdala Responsivity to Masked Fearful Eye Whites. *Science*, *306*(5704), 2061–2061. http://doi.org/10.1126/science.1103617

Wiens, S. (2008). Concepts of visual consciousness and their measurement. *Advances in Cognitive Psychology*, *3*(1–2), 349–359. http://doi.org/10.2478/v10053-008-0035-y

Wiens, S., & Öhman, A. (2007). Probing unconscious emotional processes: On becoming a successful masketeer. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 65–90). New York, NY: Oxford University Press.

Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *24*(12), 2898–2904. http://doi.org/10.1523/JNEUROSCI.4977-03.2004

Footnotes

[1] Notably, not all authors have claimed that observers were completely unaware of the masked stimuli, and have instead claimed that awareness has been “restricted” (e.g. Carlson & Reinke, 2008; Mogg & Bradley, 1999). Nonetheless, it remains a matter of contention, with theoretical importance, to determine whether emotionally salient stimuli attract attention under genuine conditions of unawareness.