Assessing the influence of emotional expressions on perceptual sensitivity to faces overcoming interocular suppression

Renzo C. Lanfranco[1,2*], Hugh Rabagliati[1], David Carmel[1,3*]

[1]Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom; [2]Department of Neuroscience, Karolinska Institute, Stockholm, Sweden; [3]School of Psychology, Victoria University of Wellington, Wellington, New Zealand. *Corresponding authors: Dr. Renzo Lanfranco, Department of Neuroscience, Karolinska Institutet, Biomedicum D4, Solnavägen 9, Stockholm 171 65, Sweden. Email: renzo.lanfranco@ki.se, and Dr. David Carmel, School of Psychology, EA 618, Easterfield Building, Kelburn Parade, Wellington 6012, New Zealand. Email: david.carmel@vuw.ac.nz

Data and materials are publicly available on the Open Science Framework (Lanfranco, 2022, February 1) and can be accessed at https://osf.io/m83qv/

ABSTRACT

Detecting faces and identifying their emotional expressions are essential for social interaction. The importance of expressions has prompted suggestions that some emotionally-relevant facial features may be processed unconsciously, and it has been further suggested that this unconscious processing yields preferential access to awareness. Evidence for such preferential access has predominantly come from reaction times in the Breaking Continuous Flash Suppression (bCFS) paradigm, which measures how long it takes different stimuli to overcome interocular suppression. For instance, it has been claimed that fearful expressions break through suppression faster than neutral expressions. However, in the bCFS procedure, observers can decide how much information they receive before committing to a report, so although their responses may reflect differential detection sensitivity, they may also be influenced by differences in decision criteria, stimulus identification, and response production processes. Here, we employ a procedure that directly measures sensitivity for both face detection and identification of facial expressions, using predefined exposure durations. We apply diverse psychophysical approaches – forced choice localisation, presence/absence detection, and staircase-based threshold measurement; across six experiments, we find that emotional expressions do not alter detection sensitivity to faces as they break through CFS. Our findings constrain the possible mechanisms underlying previous findings: faster reporting of emotional expressions' breakthrough into awareness is unlikely to be due to the presence of emotion affecting perceptual sensitivity; the source of such effects is likely to reside in one of the many other processes that influence response times.

Keywords: Emotional expressions; continuous flash suppression; awareness; signal detection theory; face processing.

INTRODUCTION

Our facility at perceiving faces and interpreting their expressions is a critical component of social cognition, allowing us to quickly identify individuals, attribute mental states, and guess intentions (Grill-Spector et al., 2017; Jack & Schyns, 2015). In this context, the processing of facial expressions is particularly key, and several studies suggest that those expressions are processed in special or distinct ways. For example, angry expressions are detected faster than non-threatening ones (Fox et al., 2000; Krysko & Rutherford, 2009). Furthermore, the presence of an angry or fearful face alone can enhance the detection of a target stimulus shown nearby (Fox, 2002; Wilson & MacLeod, 2003). Thus, there is evidence to suggest that the processing of facial expressions is, in some sense, special (Farah et al., 1998; Oruc et al., 2019).

The importance and specialised processing of emotional expressions has prompted suggestions that emotional expressions may be processed unconsciously; these suggestions are supported by neural evidence for facial expression processing under various kinds of suppression from awareness (Almeida et al., 2013; Troiani et al., 2014; Whalen et al., 1998; Williams et al., 2004). It has been further claimed, however, that some emotional expressions enjoy preferential access to awareness compared to others; this has been supported by behavioural findings showing that faces with strong (typically negative) emotional expressions are more likely to break through the masking effects of suppression techniques (Alpers & Gerdes, 2007; Carlson & Reinke, 2008; Hedger et al., 2015; Sterzer et al., 2011; E. Yang et al., 2007). The most prominent line of evidence for this consists of reports that when emotional faces are masked from awareness using Continuous Flash Suppression (CFS), a strong form of binocular rivalry, they break through that suppression more quickly than unemotional faces, as indicated by faster response times for reporting them; this method has become known as the breaking CFS, or bCFS, paradigm (Capitão et al., 2014; Hedger et al., 2015; Stein & Sterzer, 2012; E. Yang et al., 2007; Y.-H. Yang & Yeh, 2018). Findings like this have contributed to strong claims about the nature of unconscious emotion processing – and unconscious processing in general – like the idea that most cognitive functions may not require awareness to occur (Hassin, 2013; but see also Hesselmann & Moors, 2015) and that some disfunctions in the processing of emotional stimuli in psychiatric disorders may therefore occur in absence of awareness (Capitão et al., 2014; Jusyte et al., 2015; Sterzer et al., 2011; Sylvers et al., 2011).

However, as we describe below, there are reasons to wonder what specific functions are implicated in findings of faster breakthrough. For example, some unconscious emotional processing results have been attributed to processing of low-level features rather than emotional content (Gray et al., 2013; Stein et al., 2018), and findings of emotion processing under CFS have been inconsistent (e.g., Schlossmacher et al., 2017), precluding consensus on underlying mechanisms. We describe these issues in more detail below, and also address the different processes that might underlie responses in the main paradigm used to generate these data, the bCFS technique (for reviews, see Lanfranco et al., 2023; Stein, 2019). These considerations raise the question of which of the many processes that influence response times in perceptual reports is affected by emotional expressions; elucidating this issue would constrain the conclusions that can be drawn from such studies.

Studies using bCFS have tacitly assumed that an observer's reported awareness of a face is accompanied by the emergence of an ability to make objectively veridical perceptual judgments, and that their results thus indicate that this ability arises faster for faces with an emotional expression. Indeed, such studies often ask participants to report where on the screen the stimulus appeared (a spatial localisation task), as a way of verifying that they were aware of the stimulus (e.g., Yang et al., 2007; Capitão et al., 2014; Yang & Yeh, 2018). However, the dependent measure in bCFS studies is response times, which are not a measure of perceptual sensitivity. In this paper, we report a set of studies that aimed to assess whether the presence of an emotional expression indeed has an effect on measures of perceptual detection sensitivity. Our studies built on a new variant of the bCFS paradigm (Lanfranco, Stein, et al., 2022) that enables collection of signal-detection theoretic measures of sensitivity; they used a wide variety of stimuli, to address concerns about generalisation; and they had high statistical power, to address concerns about replicability. Under these rigorously controlled conditions, we found that when overcoming suppression, emotional expressions do not confer a perceptual sensitivity advantage.

The Breaking Continuous Flash Suppression (bCFS) paradigm: inconsistencies in findings and interpretations

While unconscious processing of emotional expressions has been evaluated using a number of different experimental techniques, the most prominent evidence that such expressions enjoy preferential access to awareness has come from the bCFS paradigm (Lanfranco et al., 2023). In CFS, stimuli are masked by presenting them to only one eye, while a continuous stream of high-contrast Mondrian-like masks is shown to the other eye, dominating awareness. Although this is a powerful suppression method, the suppressed stimulus eventually breaks through it, briefly becoming visible. In the bCFS method, participants are asked to provide a response as soon as the target stimulus breaks through suppression into awareness. Their latency to do this has been taken as an index of unconscious processing which, in turn, leads to awareness, with the underlying assumption that faster breakthrough times indicate faster, more efficient, or higher priority unconscious processing. For example, outside of the domain of emotional expressions, Jiang et al. (2007) found that upright faces break through suppression faster than inverted faces, suggesting that holistic face processing may occur unconsciously and promote access to awareness (Akechi et al., 2015; Gayet & Stein, 2017; Kobylka et al., 2017; Lanfranco, Stein, et al., 2022; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011).

Using this method, E. Yang et al. (2007) reported that fearful expressions broke through suppression faster than neutral expressions, and thus suggested that emotional information can be extracted unconsciously. This finding has been highly influential (for a review and meta-analysis, see Hedger et al., 2016), and the basic experimental procedure and finding have been replicated several times. For example, Y.-H. Yang and Yeh (2018) replicated the finding that fearful expressions broke through faster than neutral ones, as did Gray et al. (2013) and Hedger et al. (2015, but see below for their alternative explanation). Capitão et al. (2014) found that fearful expressions break through faster than happy expressions, and moreover found that a participant's level of anxiety modulates this difference. Similarly, Sterzer et al. (2011) found that fearful expressions break through faster than happy and sad faces in typical participants, but that these differences are less clear in participants diagnosed with depression (but see also Münkler et al., 2015). Thus, these studies are consistent with the claim that emotional processing of faces can occur unconsciously, influencing access to awareness.

However, this interpretation is controversial because not every study has replicated the original findings. For example, while both Capitão et al. (2014) and Sterzer et al. (2011) found that fearful expressions broke through faster than happy expressions, they did not find that fearful expressions break through faster than neutral ones, contrary to E. Yang et al. (2007). Moreover, in Stein and Sterzer's (2012) stimulus set (which used schematic rather than photographed faces), happy expressions in fact broke through suppression faster than neutral, angry, and sad expressions, which is surprising given the original findings. Indeed, Hedger et al's. (2016) meta-analysis concluded that, while there was some evidence that fearful faces break through suppression faster, there was also strong heterogeneity in the literature, indicating that it is difficult to draw conclusions about the robustness and size of any effect.

In parallel to this, a related set of concerns have been raised as to whether the findings of faster breakthrough are actually driven by face/emotion-specific processes, or rather by lower-level properties of the stimuli, which make masking less effective for some images (Lanfranco et al., 2023; Pournaghdali & Schwartz, 2020; Stein, 2019). For example, Stein and Sterzer (2012) argued that suppression times were mainly affected by the relative curvature of features in a schematic face (e.g., between the mouth and face outline), rather than by its emotional expression. Similarly, Gray et al. (2013) have argued that fearful photographed faces break suppression faster than neutral faces because of particular low-level visual properties, such as their luminance and spatial frequency profiles. Yet these alternative explanations are themselves hard to evaluate, given the heterogeneity of findings in this literature. To the degree that it is unclear whether particular (e.g., fearful) facial expressions actually break through suppression faster, it is also unclear whether factors like spatial frequency offer a helpful alternative explanation.

Limitations of the Breaking Continuous Flash Suppression (bCFS) paradigm

We suggest that the difficulty in establishing whether emotional expressions are processed unconsciously relates to (at least) three factors. The first is sample size and statistical power: Many of the studies described above had relatively small samples sizes, on the order of 10 to 20 participants, and thus had low statistical power to detect small effects. Importantly, low power not only increases the probability of finding a false negative, but also means that when statistically significant effects are reported, they are likely to either over-estimate the studied effect (so-called magnitude errors) or mis-estimate the sign of the studied effect (so-called sign errors; Gelman & Carlin, 2014).

The second factor concerns the breadth of the stimuli used in these studies. Most experiments have used only a small number of distinct stimuli – e.g., the studies by E. Yang et al. (2007) and Gray et al. (2013) only used faces from four individuals. Under these conditions, it is impossible to reliably establish whether an effect is general or is specific to the individual faces used. Discrepant results across studies could thus be driven by idiosyncrasies in the small set of stimuli that each study used.

The final factor is perhaps the most important: Over the last few years, significant concerns have been raised about the reliability of bCFS as a method. In particular, a large number of claims that have been made using bCFS have failed to replicate, or at least have been shown to be highly sensitive to very particular analytic decisions (for examples of failures to replicate, see Biderman & Mudrik, 2018; Moors, Boelens, et al., 2016; Moors, Wagemans, van Ee, et al., 2016; Moors & Hesselmann, 2018; Rabagliati et al., 2018). Although there are multiple possible reasons that a study might fail to replicate, one common concern is that the dependent measure used in bCFS studies – participants' response times (RTs) for reporting whether a stimulus has been seen – can be affected by a number of factors (Gayet et al., 2014; E. Yang et al., 2014): bCFS studies assume that all else being equal, participants will simply report each stimulus as soon as they become aware of it (and, as noted above, able to make objectively veridical judgments about it, indicating perceptual sensitivity); under this assumption, any effects would reflect differences in how long it takes to become aware of (and perceptually sensitive to) different stimuli. However, RTs are a measure of overall processing, encompassing the many processes – from stimulus encoding to response production – that go into producing a *speeded* (not just correct) response. For example, while it may seem reasonable to interpret bCFS findings as suggesting that perceptual sensitivity arises faster for emotional expressions that for neutral expressions, RT measurement cannot lead to this conclusion because RTs are not a measure of sensitivity.

In bCFS studies, specifically, there are several other processes that may contribute to RT differences. First, differences in reported breakthrough times could also be caused by differences in participants' response criteria, i.e., their willingness to report a signal. Notably, breakthrough from CFS unfolds over a brief – but not immediate – time period; typically, a small part of the suppressed stimulus breaks through first, and visibility then expands to the rest of the stimulus. The amount of breakthrough that a participant requires in order to commit to reporting the stimulus may vary systematically by condition. For example, participants may possess the same perceptual sensitivity to each emotional expression category, but may be more willing (or require the accumulation of less information) to report that they have seen a fearful or angry expression than a neutral expression, and thus report the former faster, even if both stimuli take the same time to break through suppression.

Past studies have tried to control for criterion effects. In one approach, researchers have included a non-rivalrous control condition where the target stimuli are shown on top of the flashing CFS masks (Akechi et al., 2014; Costello et al., 2009; Jiang et al., 2007; Li & Li, 2015; Madipakkam et al., 2015; Mudrik et al., 2011; Paffen et al., 2018; Stein & Sterzer, 2012; Zhou et al., 2010), with the assumption that if non-rivalrous conditions emulate all processes that are not CFS-specific but that contribute to RTs, any differences found in the rivalrous condition (compared to the non-rivalrous condition) should index the process that leads to breakthrough. However, because target stimuli in non-rivalrous control conditions are more easily discernible from the mask (Stein, Hebart, et al., 2011), non-rivalrous conditions may not reproduce criterion effects present in CFS conditions.

In another approach, researchers have asked participants to perform an orthogonal task such as reporting a stimulus feature that is irrelevant to the experimental manipulation (e.g., Gayet et al., 2016; Salomon et al., 2013), with the assumption that if participants are not required to identify the experimentally critical but task-irrelevant feature, their RTs will reflect processes that are not influenced by criterion effects related to that feature. However, this raises a second concern about a process that may contribute to RTs in bCFS: Participants may still perceive the manipulated task-irrelevant feature, which may influence their decision to commit to a response regardless of whether it has anything to do with the task. Most bCFS studies involve asking participants to report the presence or location of the target stimulus as soon as it breaks through suppression, ignoring its identity (e.g., stimulus category). To use such tasks, one must assume that neither reporting the presence nor the location of a stimulus should be influenced by identification or classification of the stimulus-category, let alone any more complex recognition processes. But in bCFS tasks, participants have control over the stimuli's exposure duration, so there can be no certainty that they are indeed limiting themselves to detecting or localising the stimulus, with no involvement of identification processes. Indeed, past claims that emotional expressions are processed automatically (e.g., E. Yang et al., 2007; Y.-H. Yang & Yeh, 2018) suggest that it may be impossible to avoid engaging such processes, and emotional expression processes are themselves known to be prone to biases – for example, experiments using briefly-presented, backward-masked faces (Mihalache et al., 2021) have demonstrated a bias toward reporting anger in a task that explicitly required participants to categorise a face's specific emotional expression. RTs in bCFS localisation tasks may therefore be confounded by identification-related decision biases.

A third concern arises from the inevitable involvement of motor activity production in RT measures. Studies using bCFS assume that post-perceptual motor processes – including the decision to make a specific response (e.g., to press a particular key indicating stimulus location), preparation of the relevant motor plan, and motor activity production – all unfold at an equal rate for different stimulus categories. However, emotional expressions affect arousal, as indicated by both physiological (Kreibig, 2010; Lang et al., 1993) and neural measures (Balconi & Pozzoli, 2003; Junghöfer et al., 2001). Arousal may in turn affect motor activity (Fredrikson et al., 1998; Kreibig, 2010), increasing the speed of post-perceptual response preparation and execution.

The influences listed above (and potentially others) are not mutually exclusive, and may each contribute to the RTs measured in bCFS tasks. Furthermore, the relative contribution of each to bCFS effects may differ for different stimulus manipulations (e.g., manipulations of emotional expressions, gaze direction, face orientation, etc.). Discerning the contribution of various factors to reported stimulus visibility when overcoming CFS is thus a formidable task; not only must it isolate or at least distinguish distinct processes, but this must be done separately for different manipulations.

In the case of emotional expressions and the claim that they break through CFS faster, an appropriate first step is to assess whether the presence of emotion affects perceptual sensitivity, using methods that can distinguish sensitivity from, or are not susceptible to, effects on response criteria and motor activity. Such an investigation should also disentangle detection from identification, especially since less information is required to detect a stimulus than to identify its nature (Kobylka et al., 2017). We note that identifying the nature of a stimulus is itself a multifarious process: in most of the experiments reported here, we define "identification" as recognising that a face has an emotional expression; however, it has also been shown that identifying that a face is emotional is a separate process from recognising its specific emotion (Sweeny et al., 2013). We return to this distinction in the General Discussion.

In recent work, we have developed a variant on bCFS that is non-speeded, and specifically disentangles perceptual sensitivity from decision criterion and detection from identification (see Lanfranco, Stein, et al., 2022). This variant uses the method of constant stimuli: Participants do not decide for themselves how much evidence to accumulate before committing to a response, but instead view stimuli that are CFS-masked for predetermined amounts of time and make reports after stimulus offset in a non-speeded manner. This neutralises the possible influence of motor activity processes on RTs (because response speed is irrelevant to the dependent measures), and enables Signal Detection Theoretic (SDT) analyses of the data (Wickens, 2001; Winer & Snodgrass, 2015), allowing us to directly measure participants' perceptual sensitivity to the presented stimuli.

The present study

Here, we report a stringent exploration of whether different emotional expressions break through suppression faster than neutral expressions, as indicated by measures of perceptual sensitivity. We designed our experiments to counter the concerns raised above. Specifically, we used high-powered samples with larger numbers of participants than prior work, used a larger set of face stimuli, and used our new method of constant stimuli to either prevent (Experiments 1A-2B and 4) or directly assess (Experiment 3) criterion effects.

In our task, participants saw faces that were masked by CFS for a range of predetermined exposure durations. These durations elicited performance ranging from chance to high sensitivity on two tasks: a 2-alternative forced-choice (2AFC) decision about where on the screen the face had been located (left or right of a fixation cross) and discrimination of its expression (emotional or non-emotional). We varied the emotional expressions of the faces and, in some experiments, the orientation of the faces (upright or inverted). Participants provided a non-speeded response after each stimulus presentation. From these responses, we determined each participant's sensitivity to both the face's location (a measure of stimulus detection) and its emotional expression (a measure of category identification). If emotional expressions break through suppression faster than non-emotional expressions, then location sensitivity for the former should be greater than for the latter.

Crucially, forced-choice tasks like our localisation task are immune to detection criterion effects (Macmillan & Creelman, 1990; Peters et al., 2016; Stanislaw & Todorov, 1999): they require a decision on which of two sources of information contains the stimulus, and the stimulus is definitely present in one of them on each trial; the observer is forced to say where it appeared, and thus cannot be liberal or conservative about whether it appeared at all. In such tasks, the SDT criterion measure is thus not an indication of detection criterion, but rather a measure of response bias for one of the two possible choices – in our case, a bias for reporting left or right (as each response option is equally likely to be correct, this is orthogonal to stimulus presence). Although such biases can occur, they are irrelevant to willingness to report a face; nonetheless, we measure and report them here because they provide an indication of the extent to which participants tend to choose a consistent response when they do not know the correct answer.

Even if differences in sensitivity were to be found, this would not necessarily mean that differential criteria do not contribute to bCFS findings. One of our experiments (Experiment 3) addressed this possibility by only presenting faces on half of the trials, and replacing the forced-choice localisation with a presence/absence detection task. In this task, the SDT criterion measure indicates the extent to which an observer's willingness to report the stimulus is conservative or liberal.

To foreshadow our findings, across the six experiments reported below, we find no consistent evidence that emotional expressions affect perceptual sensitivity as faces overcome CFS. All of our data and materials are publicly available on the Open Science Framework (Lanfranco, 2021, February 1; https://osf.io/m83qv/).

EXPERIMENT 1A

In this experiment, participants saw three types of masked faces: angry, happy, and neutral. If threatening expressions such as angry expressions genuinely break suppression faster (e.g., see Almeida et al., 2013), then location sensitivity to angry faces should be greater than either happy or neutral expressions. We would expect this better sensitivity to be evident across most exposure durations, but it may also be found at a more restricted range of durations (e.g., not at the extreme ends of our range of durations, where participants would be either close to chance or close to ceiling, but rather between 2 and 3 s, which is the typical breakthrough time seen in prior bCFS studies; Gray et al., 2013; E. Yang et al., 2007).

METHOD

Participants

Thirty-four University of Edinburgh students provided informed consent and were paid £14 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Two participants were excluded from analysis (see below): the remaining 32 participants (17 female; 2 left-handed) had a mean age of 23.5 [SD_{age} = 3.8]. All the studies reported here were approved by the Psychology Research Ethics Committee of the University of Edinburgh. A new sample of participants was recruited for each experiment. All participants provided informed consent in accordance with the Declaration of Helsinki.

Past bCFS studies that have found statistically significant effects of emotional expression on breakthrough times employed around 16 participants per experiment (e.g., E. Yang et al., 2007). We doubled this number to increase power and allow counterbalancing of experimental blocks with a multiple of 8 (see Procedure). A power analysis conducted using G*Power 3.1.9.7 (Faul et al., 2009) to test for differences between conditions in a repeated-measures ANOVA, with small to medium effect size ($\eta p^2 = .04$) and alpha of .05, aiming to a statistical power of 95%, determined that a sample of 19 participants would be required. If a non-sphericity correction ε of .5 were to be added – as reported below, a few tests violated the sphericity assumption – then 29 participants would be required. This analysis supports our initial decision of aiming for 32 participants per experiment.

Stimuli and Apparatus

Stimuli were presented on a 19-inch CRT monitor in a dimly lit room, connected to a computer running Matlab 2014a (Mathworks, Inc) using the Cogent 2000 toolbox (www.vislab.ucl.ac.uk/cogent.php). A chin rest and mirror stereoscope were positioned 57 cm from the monitor, with a vertical divider splitting the display so each eye only saw half of the screen.

To maintain binocular alignment, two vertical textured vergence bars (width 1°, height 8°) appeared to the left and right of fixation in each eye (horizontal centre-to-centre distance 3.1°); thanks to the stereoscope, the two eyes' displays overlapped such that only one pair of vergence bars was perceived (Carmel et al., 2010). A black fixation cross (0.7° × 0.7°) was presented in the centre of each pair of vergence bars. Rectangular Mondrian-like masks differing in size, grey-level, rotation, and position were flashed at 10 Hz to one eye while a face stimulus was introduced to the other eye.

Stimuli were 60 human faces, comprising 20 individual identities (10 male, 10 female) which were chosen from the Karolinska Directed Emotional Faces (KDEF) database (Goeleven et al., 2008). All were Caucasian and seen from a front angle. For each individual, we included three facial expressions: angry, happy, and neutral, resulting in 20 items for each expression (using the same identities minimised identity-related feature differences across expressions). Images were cropped to show only the internal facial features and transformed to greyscale (width 3.6°, height 6.1°). Luminance was equated for all the resulting images using the Matlab SHINE toolbox (Willenbockel et al., 2010). Contrast was not normalised across stimuli; notably, recent findings have shown that doing so may artifactually inflate the perceived salience of negative expressions (Webb et al., 2020). Background colour was replaced with uniform grey, matching the screen's background colour. Stimuli were then sorted into 6 different sets for counterbalancing purposes (see below), 2 for each expression, with an equal number of male and female faces. We used published norms from the KDEF validation study (Goeleven et al., 2008) to select face stimuli that – while maintaining the requirement to have 10 female and 10 male identities, each contributing exemplars of all three expressions – were matched on the norms for perceived intensity of angry and happy sets (expression intensity: M_{angry} = 5.68 [0.93]; M_{happy} = 5.85 [0.69]; $M_{neutral}$ = 5.15 [0.41]). It was impossible to match both expression intensity and identifiability simultaneously, so we prioritised intensity matching while minimising as much as possible differences in the norms for expression identification between the three expressions (expression identification: M_{angry} = 85% [SD_{angry} = 16.87]; M_{happy} = 94.6% [6.14]; $M_{neutral}$ = 83.02% [8.36]).

Procedure

Each trial began with a fixation cross, presented binocularly at the centre of each eye's visual field between two vergence bars. After 200 ms a changing Mondrian-like mask was presented to one eye (the fixation cross remained superimposed on the CFS mask in that eye) and a face image was introduced to the other eye, either to the left or to the right side of the cross (Figure 1a). We counterbalanced which eye received the mask/stimuli across participants. Following E. Yang et al. (2007), the face's contrast increased linearly from 0% to 25% over 1 s, after which it remained stable. Stimuli were presented for one of seven predefined durations, spaced equally on a log scale (600; 900; 1350; 2025; 3038; 4557; 6836 ms), with an equal number of trials for each duration. Following piloting, these durations were anticipated to capture the full range from chance-level to near-perfect performance. After 1 s, the mask's contrast began decreasing linearly until reaching zero at 6 s (so that faces presented for 6836 ms were not masked at the end; Figure 1b). Emotional expression was blocked (70 trials/block). Each block comprised either angry and neutral faces, or happy and neutral faces. Block order was ABBA BAAB BAAB ABBA, ensuring the same average positioning of block types in the sequence, with the meaning of A and B (blocks containing angry or happy faces with their neutral counterparts) counterbalanced across participants.

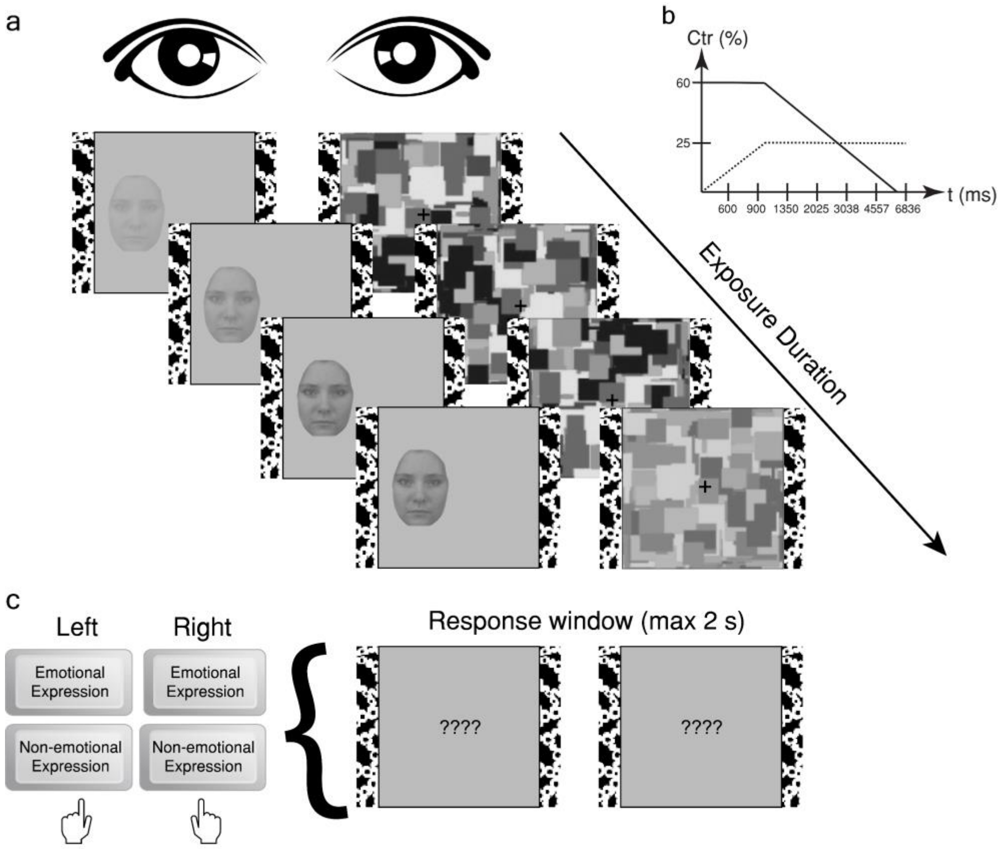

Figure 1. Schematic description of a trial in experiments 1A through 2B. (a) The duration of each display was chosen from seven equally likely exposure durations. (b) The contrast of the target image increased linearly until reaching 25% contrast at 1 s. Then, contrast remained unchanged until the end of the display. Mondrian-like patterns started at 60% contrast until reaching 1 s., after which contrast linearly decreased until reaching 0% at 6 s of exposure. (c) After stimulus offset, participants were presented with a binocular response cue composed of question marks. Participants provided a single response to indicate both on which side of the fixation the face had been shown and whether its expression was emotional or neutral.

Participants were instructed to focus on the fixation cross with both eyes open, trying to avoid blinking and eye movements. After stimulus presentation, the fixation cross was replaced by four question marks and participants were instructed to use a single keypress to report (within 2 s) both the location of the face image and whether its expression was emotional or non-emotional. To do so, they pressed one of four keys (Figure 1c); the keys indicating emotional/non-emotional expression were counterbalanced across participants. Note that this response structure dissociates stimulus detection (the localisation response) from emotion identification. After each trial, an intertrial interval (ITI) of 1000 ms, during which only the vergence bars were presented, elapsed before the next trial began. Participants were given self-terminated breaks every 70 trials and a compulsory 15-minute break halfway through the experiment. Before beginning the experiment, participants completed 60 training trials to ensure that the stereoscope was properly calibrated and that they had understood the task.

Analysis

We excluded data from one participant who failed to respond on more than 5% of trials, and one who showed chance-level accuracy at all durations, suggesting that they failed to attend to the task. Trials that received no response were treated as missing data.

We used signal detection measures to separately analyse how perceptual sensitivity and response criteria changed with trial durations, examining both judgments of where the stimuli were, and whether they expressed an emotion or not. To determine bias-independent sensitivity to face location (location d′) for each combination of duration and emotional expression, hits were defined as trials in which a face was displayed on the right and reported as being on the right, and false alarms (FAs) as trials in which a face was displayed on the left but reported as being on the right. To determine sensitivity to the presence of an emotional expression, hits were defined as trials in which an emotional face was shown and reported, and FAs as trials in which a non-emotional (neutral) face was shown but an emotional face was reported (note that for the sake of simplicity, we refer to this measure as emotion identification d′ in all experiments below. In most experiments this refers to identifying the presence of an emotional expression rather than identifying the specific emotion. The only exception to this is Experiment 2B, where participants identified whether faces were happy or fearful). For each measure, we calculated d′ by subtracting the Z-transformed FA rate from the Z-transformed hit rate. Since in SDT terms the location task is a 2-alternative forced-choice task (requiring a decision on which of two sources of information contains the signal), for this task we divided d′ by the square root of 2 (Macmillan & Creelman, 1991; Stein & Peelen, 2021; Wickens, 2001). A d′ of 0 indicates chance performance. We also calculated each participant's criterion measures (C) for both tasks, by multiplying each task's sum of Z-transformed hit and FA rates by -0.5 (Macmillan & Creelman, 1991). For the location task, positive and negative values for this measure are not a measure of criterion, but rather indicate a bias toward responding "left" and "right", respectively; but these may be idiosyncratic and cancel out across participants, so we converted the results to absolute values to assess the magnitude of response biases, independently of their direction. For the identification task, lower values indicate that the participant is more willing to report emotional expression.

For each participant, we calculated location d′, response bias, expression identification d′, and emotion identification criterion for every combination of emotional expression and exposure duration. We analysed these measures using repeated-measures analysis of variance (ANOVA); wherever Mauchly's test indicated that the sphericity assumption was violated, degrees of freedom were corrected using Greenhouse-Geisser corrections. When interactions were significant, we explored them further using post hoc Bonferroni-corrected pairwise comparisons based on the pooled variance of the ANOVA model. Where null results were of theoretical interest, we calculated Bayes factors to evaluate the strength of the evidence for the null. Our model for the null hypothesis was a standard Cauchy distribution centred on zero with rate of .707.

Frequentist (ANOVA) and Bayesian (Bayes factors) statistical analyses were performed using Jamovi (version 1.2.17; The jamovi project, 2020) and JASP (version 0.11.1; JASP Team, 2020), respectively.

RESULTS

Location sensitivity

We calculated mean location d′ scores for angry, happy, and neutral expressions. We obtained two different d′ scores for neutral expressions, as there were neutral-expression trials in blocks containing angry-expression trials and blocks containing happy-expression trials. Following a preliminary analysis showing that d′ scores for neutral expressions did not differ in the two types of blocks (see Supplementary Material 1), we collapsed neutral-expression trials across block types, making three conditions in total: angry, happy, and neutral.

Next, we entered location d′ scores into a 3 (expression: angry, happy, neutral) x 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(2.77,\ 85.74)} = 104.537, p < .001, \eta p^2 = .771$); as Figure 2a shows, location d′ scores increased from near chance to high sensitivity with increasing exposure duration. Our key question was whether sensitivity levels would differ across the different facial expressions, either overall or in interaction with exposure duration. However, we did not find a main effect of expression, meaning that the emotional expression of the faces did not significantly affect participants' sensitivity

$(F_{(1.73,\ 53.78)} = 2.422, p = .105, \eta p^2 = .072)$. The interaction between expression and exposure duration was also not significant $(F_{(7.71,\ 239)} = 0.831, p = .572, \eta p^2 = .026)$. Bayes factors indicated very strong evidence for the null hypothesis model (BF_{01} = 46.825), i.e., no effect of expression on location d′. Thus, this first analysis provided no evidence of differential sensitivity to emotional faces compared to neutral faces.

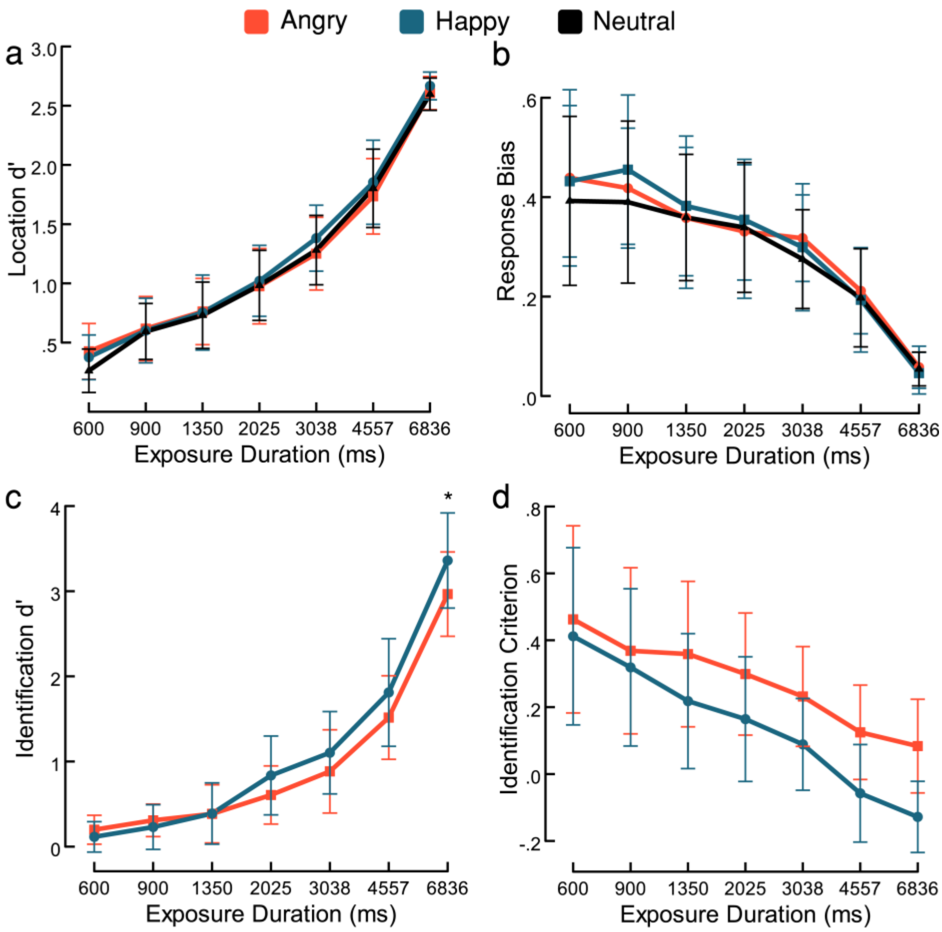

Figure 2. Results of Experiment 1A, with each dependent measure given as a function of exposure duration. (a) Location d′. (b) Response bias (absolute value). (c) Expression identification d′. (d) Expression identification criterion. (* = *p* < .05 and error bars represent 95% confidence intervals).

Response Bias

Participants' response bias for the location judgments (Figure 2b) were also analysed in a 3 (Expression: angry, happy, neutral) x 7 (exposure duration) repeated-measures ANOVA. Response bias significantly decreased with increasing exposure durations; indeed, the main effect of exposure duration was significant, $(F_{(1.97,\ 61.17)} = 14.515, p < .001, \eta p^2 = .319)$. This effect demonstrates (sensibly) that as stimulus visibility increases, observers become less likely to adopt a response bias. However, response bias was unaffected by expression ($F_{(1.88,\ 58.24)} = 1.198, p = .307, \eta p^2 = .037$). This null effect was supported by Bayes factors too ($BF_{01} = 37.885$). The interaction between expression and exposure duration did not reach significance ($F_{(6.78,\ 210.2)} = 0.358, p = .922, \eta p^2 = .011$). Bayes factors supported this null interaction ($BF_{01} > 100$).

Expression identification sensitivity

To examine whether participants' sensitivity to identifying the suppressed face's expression varied across the emotions assessed (Angry and Happy faces versus neutral), we entered identification d' scores into a 3 (Expression: angry, happy, neutral) x 7 (exposure duration) repeated-measures ANOVA. Identification d' scores increased with increasing exposure duration ($F_{(2.69,\ 83.29)} = 81.23, p < .001, \eta p^2 = .724$), (Figure 2c). In addition, we found a main effect of expression ($F_{(1,\ 31)} = 4.47, p = .043, \eta p^2 = .126$): Participants were more sensitive to happy expressions than angry expressions. The interaction between expression and exposure duration was also significant ($F_{(4.21,\ 130.51)} = 4.37, p = .002, \eta p^2 = .124$); follow-up post hoc comparisons revealed that the advantage of happy expressions over angry expressions only reached significance at the longest exposure duration, ($t(31) = -3.672,\ p = .03,\ d = -0.649$). Thus, while emotional expressions did not affect judgments of where a face was, they did affect the ability to judge what expression the face showed: there was greater sensitivity to the emotional content of happy expressions than angry expressions at longer exposure durations.

Expression decision criterion

As discussed in the Introduction, one potential explanation for why emotional expressions like anger break CFS faster, is that they enjoy a more liberal identification criterion. Under the conditions of this study, however, a 3 (Expression: angry, happy, neutral) x 7 (exposure duration) repeated-measures ANOVA showed that participants exhibited a significantly more liberal criterion for happy expressions than for angry expressions, as demonstrated by the main effect of expression ($M_{\text{angry}} = 0.276, SD = 0.564,$ [95% $CI = 0.115, 0.436$]; $M_{\text{happy}} = 0.145, SD = 0.549,$ [$-0.001, 0.291$]; $F_{(1,\ 31)} = 19.80, p < .001, \eta p^2 = .390$), (Figure 2d). In addition, participants' willingness to report an emotional expression (indexed by lower criterion scores) increased with increasing exposure durations ($F_{(1.65,\ 51.03)} = 7.66, p = .002, \eta p^2 = .198$). The interaction did not reach significance ($F_{(3.95,\ 122.49)} = 1.71, p = .153, \eta p^2 = .052;\ BF_{01} = 63.502$).

DISCUSSION

Experiment 1A tested whether emotional expressions enjoy a sensitivity advantage over neutral expressions in breaking through CFS. We used a new variant of the bCFS paradigm, which allowed us to disentangle detection from identification and directly measure sensitivity. Our data, however, show no evidence that angry expressions break through CFS any faster (as indexed by location d') than either neutral or happy expressions.

Surprisingly, although expressions had no effect on location sensitivity (indicating that they did not affect detection), we found that participants exhibited better sensitivity to identifying the presence of emotion in happy expressions than angry expressions, i.e., the former were easier to distinguish from their neutral counterparts than the latter. Based on past bCFS findings, one may have expected either better identification sensitivity, a more liberal identification criterion, or both for angry expressions over happy ones. However, it is important to note that the identification advantage of happy expressions over angry ones only reached significance at the longest exposure duration, once CFS masks had disappeared. Hence, this effect could have been driven by inherent differences between happy and angry expressions in the KDEF stimulus set; as noted in the Method section, we used published norms (Goeleven et al., 2008) to select stimuli that equated the intensity of angry and happy expressions while minimising differences in expressions' identifiability. Because perfect simultaneous matching of both intensity and identifiability was impossible, happy faces were slightly more identifiable, and this may account for the identification d' advantage and the more liberal criterion seen for them. With this in mind, we note that most previous bCFS studies did not attempt stimulus matching for both intensity and identifiability of expressions; our finding of better identification d' and a more liberal criterion for reporting the presence of emotion in happy over angry expressions reinforces the idea that identification-related processes could affect detection response times, highlighting the importance of disentangling detection from identification.

Our null finding regarding location sensitivity suggests that neither emotional expression enjoys a detection sensitivity advantage. However, it is also possible that the particular emotional expressions used here – happy and angry – may not be ideal for evoking differential sensitivity. In particular, the meta-analysis by Hedger et al. (2016) found that fearful expressions presented a more robust and consistent effect in bCFS studies searching for differential breakthrough times for different emotions. Therefore, in our next experiment we used fearful expressions instead of angry expressions.

EXPERIMENT 1B

Experiment 1B was the same as Experiment 1A, except that we replaced the angry expressions with fearful ones, and tested whether they break through suppression faster (as indicated by perceptual sensitivity measures) than happy and neutral expressions.

METHODS

Participants

A new sample of thirty-two University of Edinburgh students (19 female; 3 left-handed), with a mean age of 21.4 [SD = 4.2], participated for a payment of £14.

Stimuli and Procedure

The methods were the same as in Experiment 1A, except that fearful faces were employed instead of angry faces. These images were again selected from the KDEF database, as in Experiment 1A, we matched the intensity of fearful and happy expressions (M_{fearful} = 5.82; [0.61]; M_{happy} = 6.13 [0.84]; M_{neutral} = 4.85 [0.66]), while minimising differences in expression identification (M_{fearful} = 71%; [7.44]; M_{happy} = 96.09% [4.49]; M_{neutral} = 65.1% [18.25]).

RESULTS

Our findings broadly replicated Experiment 1A. First, as in Experiment 1A, we collapsed neutral-expression trials into one category (Supplementary Material 2). Next, in the main analyses, we entered each of the dependent measures into 3 (expression: fearful, happy, neutral) x 7 (exposure durations) repeated-measures ANOVAs. For location d' scores, we found that location sensitivity increased with increasing exposure duration (Figure 3a, $F_{(2.09,\ 64.82)} = 111.961, p < .001, \eta p^2 = .783$), and there was no effect of expression ($F_{(1.84,\ 57.15)} = 0.103, p = .887, \eta p^2 = .003;\ \mathrm{BF}_{01} = 58.6$) nor an interaction between expression and exposure duration ($F_{(6.81,\ 211.26)} = 1.042, p = .402, \eta p^2 = .033, \mathrm{BF}_{01} > 100$). Thus, as before, emotional (and in particular, fearful) expressions did not confer a location-sensitivity advantage over non-emotional expressions in overcoming CFS.

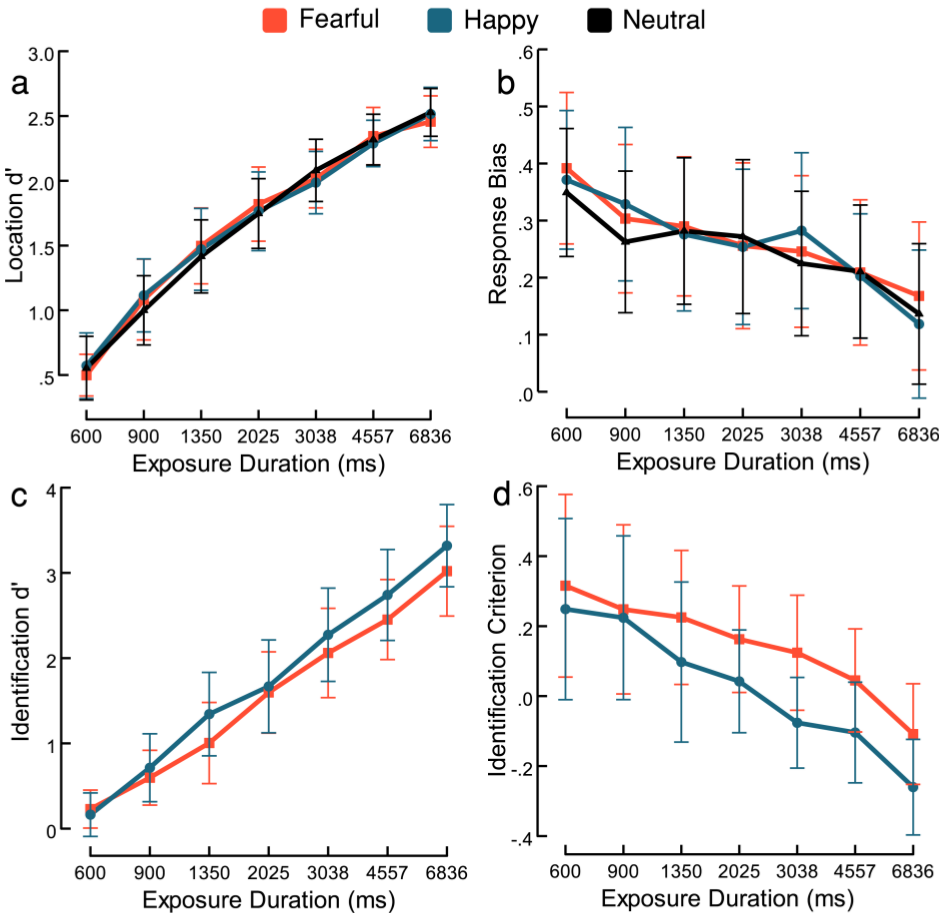

Figure 3. Results of Experiment 1B. (a) Location d'. (b) Response bias. (c) Expression identification d'. (d) Expression identification criterion. Error bars represent 95% CI.

Second, response bias significantly decreased with increasing exposure duration ($F_{(2.84,\ 88.13)} = 8.001, p < .001, \eta p^2 = .205$), (Figure 3b), but as in Experiment 1A, this decrease was not modulated by expression ($F_{(1.74,\ 53.90)} = 0.562, p = .550, \eta p^2 = .018;\ BF_{01} = 35.461$). The interaction did not reach significance either ($F_{(8.19,\ 253.89)} = 0.753, p = .647, \eta p^2 = .024;\ BF_{01} > 100$).

Third, identification d' scores increased with increasing exposure duration ($F_{(2.41,\ 74.7)} = 76.89, p < .001, \eta p^2 = .713$), (Figure 3c), and similar to Experiment 1A we found a main effect of expression ($F_{(1,\ 31)} = 9.86, p = .004, \eta p^2 = .241$), with an advantage of happy expressions over fearful expressions. The interaction between expression and exposure duration also reached significance

($F_{(5.32,\ 165.01)} = 2.34, p = .04, \eta p^2 = .07$). Inspection of Figure 3c suggests that the identification effect was larger at longer durations (as in Experiment 1A), although Bonferroni-corrected post hoc comparisons did not reveal significant differences between expressions at individual exposure durations.

Finally, there was a main effect of exposure duration on expression identification criteria ($F_{(1.44,\ 44.58)} = 7.54, p = .004, \eta p^2 = .196$), (Figure 3d), and a more liberal identification criterion for happy expressions ($M_{\text{happy}} = 0.145, SD = 0.546,$ [95% $CI = -0.123, 0.172$]) than for fearful expressions ($M_{\text{fearful}} = 0.276, SD = 0.539, [-0.005, 0.294]; F_{(1,\ 31)} = 16.08, p < .001, \eta p^2 = .342$). Thus, overall, Experiment 1B broadly replicated the findings of Experiment 1A, but using fearful rather than angry faces.

DISCUSSION

Experiment 1B, like Experiment 1A, did not find an advantage in location sensitivity for emotional expressions, this time using fearful instead of angry expressions. Importantly, the fact that neither angry nor fearful expressions enjoyed a sensitivity advantage over happy or neutral expressions suggests that negative threat-related faces do not enjoy perceptual priority over non-threat-related faces in overcoming CFS.

Similar to Experiment 1A, Experiment 1B found better identification sensitivity for positive (happy) expressions than negative (here, fearful) expressions, and a more liberal decision criterion for happy expressions than fearful ones. As noted in the Discussion of Experiment 1A, these results may be due to idiosyncrasies of our stimulus set; importantly, however, they are consistent with the possibility that prior bCFS results may be explained by similar idiosyncrasies, e.g., that identification processes may have influenced response times.

Whilst the results of experiments 1A and 1B are consistent with each other, they do not rule out the possibility that our procedure may not be sufficiently sensitive to capture effects of emotional expression on detection (location d'). In order to verify that the method is capable of revealing effects of facial attributes, we designed Experiment 2A to allow us to simultaneously test for an effect of emotional expression while also attempting to capture a different face-related effect that has been replicated several times in the bCFS literature: the face-inversion effect (FIE). Multiple bCFS studies have shown that upright faces have shorter breakthrough times than inverted faces (Akechi et al., 2015; Gayet

& Stein, 2017; Jiang et al., 2007; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011; Zhou et al., 2010; see also Lanfranco, Stein, et al. (2022) for previous evidence that this method can capture FIEs). Thus, we reasoned that combining a test of emotional expression with a test of face inversion can help us to calibrate the sensitivity of our method.

EXPERIMENT 2A

In this experiment, we only used fearful and neutral facial expressions, but presented each in both upright and inverted orientations. If our method is indeed sensitive to the properties of these stimuli, then we expect to see inversion affecting location sensitivity.

METHODS

Participants

We recruited a new sample of thirty-two University of Edinburgh students (17 female; 4 left-handed; mean age of 23.1[3.9]), who were paid £14 for their participation.

Stimuli and Procedure

These were similar to Experiment 1B, but only using the fearful and neutral faces. We selected all the stimuli from the KDEF database again to ensure they could be matched on the criteria described in the previous experiments (expression identification: M_{fearful} = 73.1%; [9.82]; M_{neutral} = 73.06% [13.34]; expression intensity: M_{fearful} = 5.9; [0.78]; M_{neutral} = 4.98 [0.54]). Faces were presented in upright and inverted orientations. Orientation was blocked (70 trials/block) with block order counterbalanced across participants as in Experiment 1B. Thus, participants went through upright and inverted blocks, each containing images of both fearful and neutral faces.

RESULTS

We submitted location d' and location response-bias scores to 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVAs. Similar to the previous experiments, location d' scores increased with increasing exposure duration (Figure 4a, $F_{(2.05,\ 63.47)} = 185.273, p < .001, \eta p^2 = .857$).

Furthermore, emotional expression again had no effect, as in Experiments 1A and 1B ($F_{(1,\ 31)} = 2.03, p = .164, \eta p^2 = .061$); this null effect was supported by Bayes factor analysis ($BF_{01} = 10.551$). In contrast, orientation did affect location d' ($F_{(1,\ 31)} = 50.409, p < .001, \eta p^2 = .619$), with an advantage for upright faces over inverted faces. In addition, the interaction between orientation and exposure duration was significant ($F_{(4.53,\ 140.29)} = 3.576, p = .006, \eta p^2 = .103$). There was a pattern of better location sensitivity for upright over inverted faces at all durations except the extreme ends of the range, with Bonferroni-corrected post hoc comparisons showing significant differences between upright and inverted faces at 1350 ($t(209) = 5.182,\ p < .001,\ d = 0.916$) and 3038 ms of exposure ($t(209) = 5.596,\ p < .001,\ d = 0.989$).

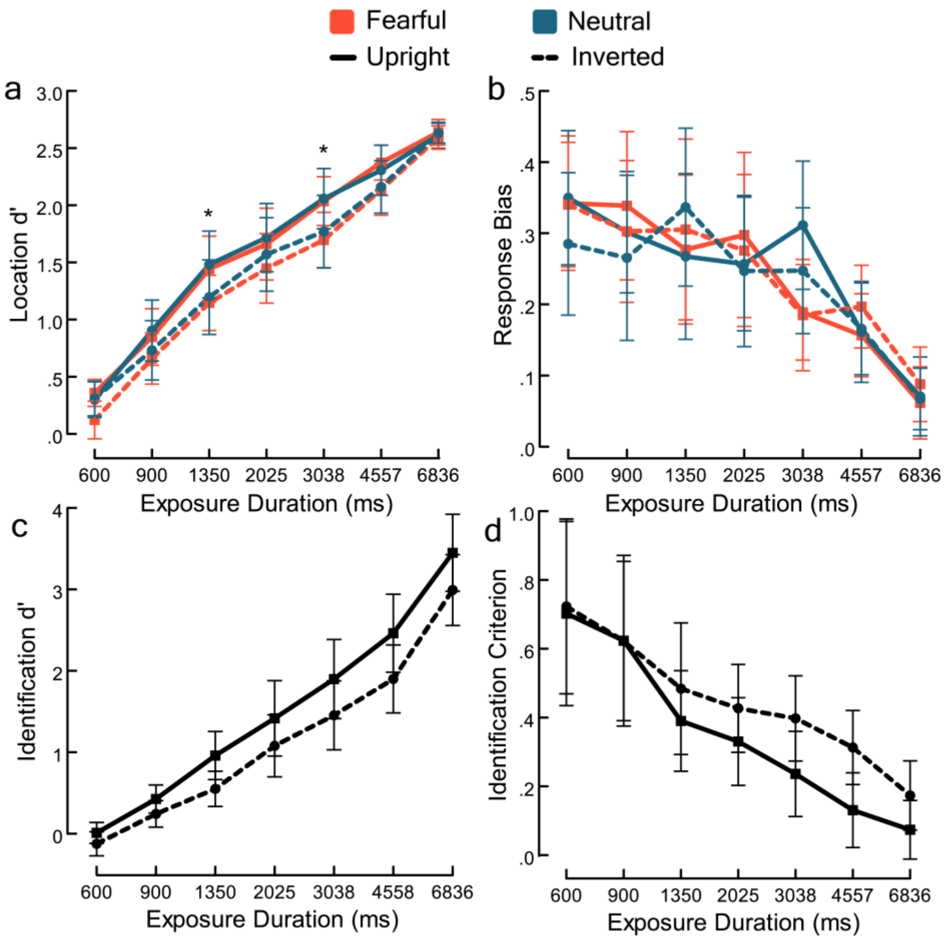

Figure 4. Results of Experiment 2A. (a) Location d'. (b) Response bias. (c) Expression identification d'. (d) Expression identification criterion. (* = *p* < .05 and error bars represent 95% confidence intervals).

Similar to Experiments 1A and 1B, location response bias significantly decreased with increasing exposure durations ($F_{(2.82,\ 87.42)} = 16.202, p < .001, \eta p^2 = .343$), (Figure 4b). Neither an effect of expression ($F_{(1,\ 31)} = 0.025, p = .874, \eta p^2 = .001; BF_{01} = 13.638$) nor of orientation ($F_{(1,\ 31)} = 0.151, p = .7, \eta p^2 = .005;\ BF_{01} = 12.261$) was found. Unexpectedly, we found a significant interaction between expression and exposure duration ($F_{(4.59,\ 142.38)} = 2.426, p = .043, \eta p^2 = .073$). However, Bonferroni-corrected pairwise comparisons did not reveal significant differences between expressions at any exposure duration. Finally, no other interaction reached significance (all *p* values > .351).

We submitted identification d' and identification criterion scores to 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVAs. Identification d' scores significantly increased with increasing exposure duration ($F_{(2.54,\ 78.69)} = 88.46, p < .001, \eta p^2 = .740$), (Figure 4c). Importantly, participants were more sensitive to the emotional expression in upright faces than in inverted faces, as indicated by higher identification d' for upright than inverted faces ($F_{(1,\ 31)} = 28.38, p < .001, \eta p^2 = .478$). Finally, we found a marginal interaction between duration and orientation ($F_{(4.37,\ 135.53)} = 2.36, p = .051, \eta p^2 = .071$); exploratory Bonferroni-corrected post-hoc comparisons revealed significant differences between upright and inverted faces at 1350 ($t(147) = 3.965, p = .037, d = 0.701$) and 4557 ms of exposure ($t(147) = 4.2, p = .019, d = 0.743$).

Finally, participants' willingness to report an emotional expression (indexed by lower criterion scores) increased with increasing exposure durations ($F_{(1.5,\ 46.57)} = 14.34, p < .001, \eta p^2 = .316$). We also found a more liberal criterion for upright faces than inverted faces ($F_{(1,\ 31)} = 6.93, p = .013, \eta p^2 = .183$). The interaction between the two factors also reached significance ($F_{(5.17,\ 160.12)} = 2.30, p = .046, \eta p^2 = .069$), but Bonferroni-corrected pairwise comparisons did not reveal significant differences at any specific duration.

DISCUSSION

Experiment 2A provided further evidence that emotional (fearful) expressions do not have a sensitivity advantage over neutral expressions in overcoming CFS. As before, we found no effect of expression on location sensitivity. This time, however, we also tested for a different effect reported in the bCFS literature – the face inversion effect – and found that indeed, upright faces had an advantage over inverted faces in overcoming suppression. This effect was present for both location sensitivity and emotion identification, demonstrating that our method is sensitive to the nature of the stimuli used.

One remaining possibility for why we did not find an effect of expression in experiment 2A is that we compared fearful expressions to neutral expressions, and neutral expressions may not be the best comparison as they might be ambiguous and thus harder to recognise. Therefore, we ran an additional experiment, in which we used happy expressions instead of neutral ones. This allowed us to compare fearful faces (clearly threat-related expressions), with happy faces (clearly non-threat-related expressions).

EXPERIMENT 2B

We used the same methods as in Experiment 2A, but replaced the neutral-expression faces with happy faces. Fearful and happy faces were again presented both in upright and inverted orientations. Unlike our other experiments, in which the identification task required participants to report whether a face was emotional or neutral, in this experiment participants were asked to identify the expression itself by categorising it as fearful or happy.

METHODS

Participants

We recruited a new sample of thirty-two University of Edinburgh students (25 female; 2 left-handed; mean age of 20.1[2.2]), who were paid £14 for their participation.

Stimuli and Procedure

The methods were identical to those of Experiment 2A, with the following exceptions: First, we employed happy rather than neutral expressions. These images were again selected from the KDEF database and matched expression intensity (M_{fearful} = 6.12; [0.79]; M_{happy} = 5.86 [0.97]) while minimising differences in expression identification (M_{fearful} = 76.38%; [5.55]; M_{happy} = 93.75% [8.71]. Second, whereas participants had previously judged whether an expression was emotional or not, in this experiment they judged whether each face showed a fearful or happy expression. For identification sensitivity analyses, a hit was defined as a trial with a fearful expression that was reported as presenting a fearful expression, whereas false alarm was defined as a trial with a happy expression that was reported as presenting a fearful expression.

RESULTS

Similar to Experiment 2A, we submitted location d′ and location response-bias scores to 2 (expression: fearful, happy) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVAs. Location d′ scores again increased with increasing exposure duration ($F_{(2.23,\ 69.16)} = 160.99, p <$

.001, $\eta p^2 = .839$), as in our previous experiments (Figure 5a). This time, though, the effect of orientation was only marginal ($F_{(1, 31)} = 3.511, p = .070, \eta p^2 = .102$). Strikingly, we found a main effect of emotional expression ($F_{(1, 31)} = 9.406, p = .004, \eta p^2 = .233$), driven by higher sensitivity to happy than fearful faces (indicating happy faces broke through suppression faster than fearful faces). This result is quite different from our other experiments, and indeed is different from most prior bCFS studies, where fearful faces were found to break through faster.

However, while the *p* value associated with this result is low, we should note that the numerical effect (Figure 5a) is extremely small. Because this effect contradicts previous experiments' results, we ran a Bayes factor analysis to test whether the data indeed supported a null effect model of emotion: the Bayes factor analysis showed very strong evidence in favour of a null effect of emotion ($BF_{01} = 8.673$), thus suggesting that this effect of emotion shown by the ANOVA may be spurious.

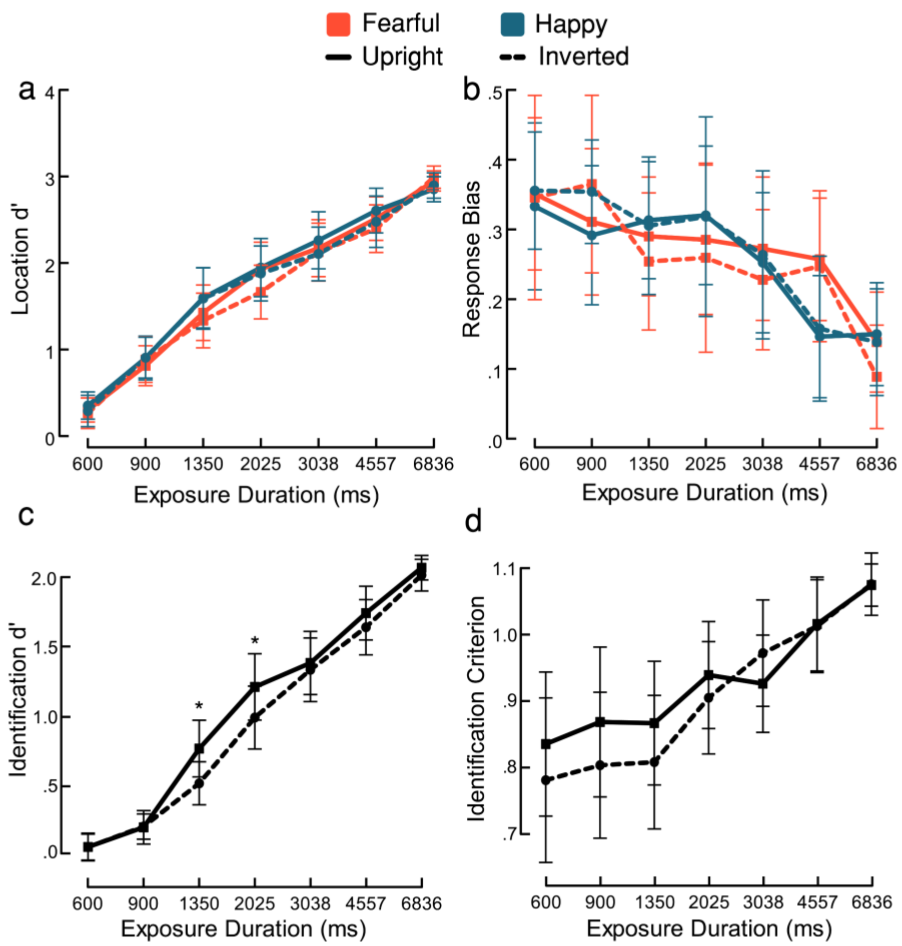

Figure 5. Results of Experiment 2B. (a) Location d'. (b) Response bias. (c) Expression identification d'. (d) Expression identification criterion. (* = *p* < .05 and error bars represent 95% confidence intervals).

Location response bias significantly decreased with increasing exposure durations ($F_{(3.26,\ 101.13)} = 8.411, p < .001, \eta p^2 = .213$), (Figure 5b). However, we did not find a main effect of expression ($F_{(1,\ 31)} = 0.001, p = .974, \eta p^2 = 0$; $BF_{01} = 13.546$), suggesting that emotional expression did not affect response bias. We did not find a main effect of orientation either ($F_{(1,\ 31)} = 0.018, p = .895, \eta p^2 = .001$; $BF_{01} = 13.662$), suggesting that face orientation did not affect response bias. No interaction reached significance (all p values > .088).

Similar to Experiment 2A, we submitted identification d′ and identification criterion scores to 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVAs. Expression identification sensitivity significantly increased with increasing exposure duration ($F_{(2.21,\ 68.49)} = 176.5, p < .001, \eta p^2 = .851$), (Figure 5c). We also found a main effect of orientation on identification ($F_{(1,\ 31)} = 12.79, p = .001, \eta p^2 = .292$), indicating greater identification sensitivity to upright than inverted faces; and a significant interaction between orientation and exposure ($F_{(5.03,\ 155.91)} = 3.93, p = .002, \eta p^2 = .113$), where post hoc Bonferroni-corrected post-hoc comparisons showed a significant advantage for upright over inverted expressions at 1350 ms ($t(147) = 4.075, p = .027, d = 0.72$) and 2025 ms of exposure ($t(147) = 3.978, p = .035, d = 0.703$). Therefore, as in Experiment 2A, participants were more sensitive to facial expression in upright faces than in inverted faces.

Finally, we found that participants showed a greater bias for reporting happy than fearful as exposure duration increased ($F_{(2.21,\ 68.37)} = 13.66, p < .001, \eta p^2 = .306$), indexed by more positive criterion scores in this experiment (Figure 5d). However, we did not find an effect of orientation on criterion ($F_{(1,\ 31)} = 2.74, p = .108, \eta p^2 = .081$). While the interaction was significant ($F_{(4.57,\ 141.52)} = 2.4, p = .045, \eta p^2 = .072$), Bonferroni-corrected pairwise comparisons did not reveal significant differences between orientations at any exposure duration.

DISCUSSION

Experiment 2B compared fearful expressions with happy expressions. It provided two surprising results. First, we found significantly greater sensitivity to happy expressions than fearful ones, suggesting the former break through suppression faster. This deviates from our previous null findings, and contradicts prior claims (e.g., Capitão et al., 2014; Hedger et al., 2015; E. Yang et al., 2007) that it is fearful faces that break through faster. However, we think this result is likely to be a false positive: the effect's magnitude was very small, and importantly, it was not corroborated by a Bayes factors analysis. Thus, we do not place strong stock in it.

Second, the face-inversion effect in this experiment was only marginal for the detection task, though it remained significant for the identification task. This may suggest that detection tasks rely less heavily on holistic processing than identification tasks do.

So far, our experiments have employed a procedure in which participants had to make two judgments in each response: First, they had to choose the location of the face out of two possible locations and, in the same response, decide whether or not the stimulus was emotional. However, bCFS studies have typically employed a detection task where participants simply had to press a key as soon as they became aware of the stimulus (e.g., Schlossmacher et al., 2017; E. Yang et al., 2007; Zhan et al., 2015); even when the task included localisation, it did not require identification of the stimulus – an additional decision, which may increase the cognitive demands of the task. To make the cognitive demands of our task more similar to those of such bCFS studies, in our next experiment we turned our procedure into a Yes-No detection task.

EXPERIMENT 3

In Experiment 3, we simply required participants to detect faces: Participants reported whether a face had been shown or not, but not where on the screen a face had appeared nor what expression it showed, making the task's cognitive demands simpler. As in Experiment 2A, we used fearful and neutral expressions presented in upright and inverted orientations. Faces were presented either on the left or on the right side of the screen as in our previous experiments, but this time they were presented on only half of the trials, and participants were simply required to respond 'Yes' (without indicating the side) if they thought a face was shown and 'No' if they thought no face was shown.

As before, if fearful expressions break through suppression faster than neutral expressions, we should find higher detection sensitivity for fearful expressions than for neutral expressions. Additionally, unlike our previous experiments, in this experiment's present/absent paradigm the decision criterion reflects willingness to report the presence of a face (liberal or conservative). Therefore, if more liberal criteria for reporting fearful expressions can partially explain past bCFS findings, then we should find evidence that compared to neutral faces, participants are more willing to report fearful faces as present.

Participants

We recruited a new sample of thirty-two University of Edinburgh students (21 female; 2 left-handed; mean age 23.2 [SD_{age} = 3.1]), who were paid £14 for their participation.

Stimuli and Procedure

We used the same stimuli as in Experiment 2A: fearful and neutral expressions in upright and inverted orientations. The total number of trials was 1120. However, half of the trials presented a face, and the other half did not present a face. In half of the face-present trials the face had a fearful expression, and in the other half it had a neutral expression. Expression was blocked (70 trials/block) with block order counterbalanced across participants. Participants pressed the up and down arrows to report whether the face was present or absent; the keys were counterbalanced across participants.

RESULTS

We measured detection sensitivity using the standard d' formula for yes/no tasks (Macmillan & Creelman, 1991; Wickens, 2001), subtracting the Z-transformed FA rate from the Z-transformed hit rate; a hit was defined as trials in which a face was displayed and reported as being present, and FAs as trials in which no face was displayed but reported as present. We submitted Detection sensitivity (d') and detection criterion scores to 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVAs. Detection sensitivity scores significantly increased with increasing exposure duration ($F_{(2.54,\ 78.67)} = 141.812, p < .001, \eta p^2 = .821; BF_{01} = 4.947$), (Figure 6a). However, we did not find a main effect of expression ($F_{(1,\ 31)} = 0.560, p = .460, \eta p^2 = .018;\ BF_{01} = 13.102$) or of orientation ($F_{(1,\ 31)} = 2.191, p = .149, \eta p^2 = .066$). No interaction reached significance (all p values > .296).

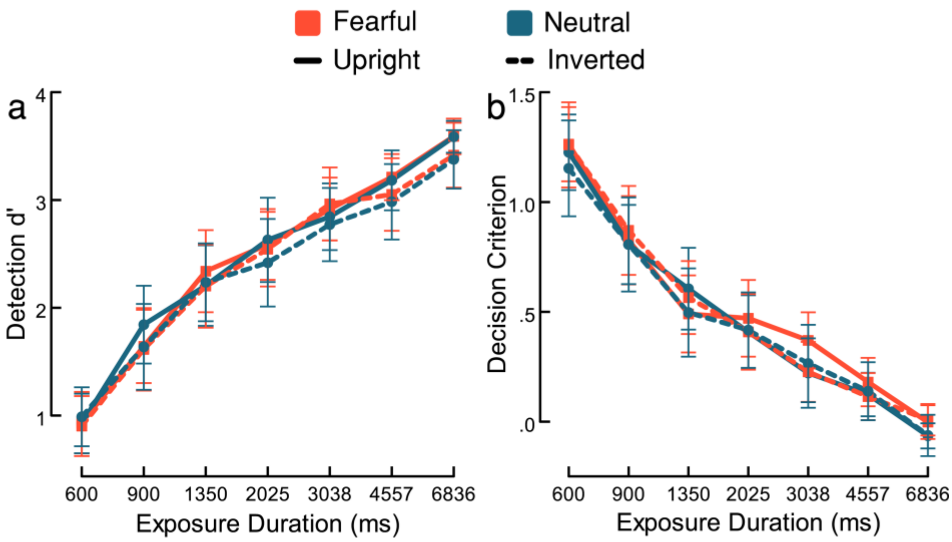

Figure 6. Results of Experiment 3. (a) Detection d'. (b) Decision criterion. Error bars represent 95% confidence intervals.

To test whether participants exhibited more liberal criteria to fearful expressions than to neutral ones, we examined whether participants' criteria varied across conditions. We found that criteria were more liberal for longer exposure durations $(F_{(2.37,\ 73.60)} = 115.271, p < .001, \eta p^2 = .788)$, (Figure 6b). We also found a main effect of expression $(F_{(1,\ 31)} = 6.554, p = .016, \eta p^2 = .175)$, which surprisingly indicated that fearful expressions were associated with more conservative criteria. However, the effect was small and probably mostly driven by two higher scores obtained at 2025 and 3028 ms of exposure. We did not find a main effect of orientation $(F_{(1,\ 31)} = 0.165, p = .688, \eta p^2 = .005)$. Finally, neither the two-way interactions (all $p > .469$) nor the three-way interaction ($p > .053$) reached significance.

DISCUSSION

Experiment 3 again found no effect of emotional expression on whether a face broke through CFS, even using a task whose requirements entailed similar cognitive demands to those of prior bCFS studies. In addition, it also found no evidence for a FIE; we return to this null finding in the General Discussion. Thus, using the method of constant stimuli across five experiments, with various emotional expressions and both forced-choice localisation and presence/absence detection tasks, we found only minimal evidence that the emotional expression of a face affected the time it took to break through suppression.

However, there is one further potential objection to our method: Although the set of predetermined exposure durations that we used encompassed the reaction times typically seen in bCFS studies, perhaps our specific durations were spaced too far apart to capture the range within which sensitivity differences between stimulus categories emerge. In bCFS, exposure durations are set by the individual subjects (with concomitant consequences for interpreting the data). It could thus be that our predetermined durations were unsuitable at the level of the individual subject. To remedy these concerns, we conducted a final study that used a staircase procedure to identify appropriately personalised presentation durations for each subject.

EXPERIMENT 4

Here we assessed how emotional expressions affected the presentation duration that elicited threshold detection. We used the same stimuli as in Experiment 2A, but employed a staircase procedure in which exposure durations changed based on participants' performance.

A similar approach was taken by Stein, Hebart, et al. (2011), who tested – using a similar task to ours – whether upright faces had shorter breakthrough thresholds than inverted faces. Stein, Hebart, et al. (2011) compared presentation of CFS-suppressed faces to a control condition in which the stimuli's contrast was ramped up from 0 to 100% in 8.5 s binocularly on top of the masks. They found shorter thresholds for upright (~1200 ms) than inverted faces (~1450 ms) in the CFS condition, but no significant threshold difference in the control condition. We developed a similar but methodologically stricter staircase procedure as we employed narrower step sizes, a higher number of trials per staircase, and more staircase repetitions per condition. In addition, we used both ascending and descending staircases, and randomly interleaved two staircases of each type (Stein, Hebart, et al., 2011, only used one staircase with an initial exposure duration based on previous experiments' results). Thus, our procedure yielded 16 staircases: 2 expressions × 2 orientations × 2 staircase types (ascending and descending) × 2 of each staircase type. Thresholds were average for each combination of expression and orientation, collapsing across the four staircases of each combination of expression and orientation, thereby providing more robust threshold estimates. Additionally, we employed 33 participants whereas Stein, Hebart, et al. (2011) had only 13.

METHODS

Participants

We recruited a new group of thirty-three University of Edinburgh students (23 female; 4 left-handed; mean age 22.7[2.8]), who were paid £14 for their participation.

Stimuli and Procedure

We used a two-stage staircase procedure. The staircases in the first stage had a wide range of exposure durations (32, from 200 to 6826 ms, on a log scale with each step 1.17 times longer than the one before). Participants moved around the staircase using a 1-up, 2-down rule – i.e., two consecutive correct responses decreased the exposure duration by one step, and one incorrect answer increased it by one step. After reaching five reversals on a staircase, participants entered its second stage. The exposure duration vector was replaced by a new vector. This was centred on that fifth reversal value, and had 12 steps on either side of it with a step-size of 50 ms (constrained not to fall below 50 or above 6826 ms). This new vector remained unchanged for the rest of the experiment.

The experiment comprised 1120 trials, which were sorted into 16 different staircases, each defined by a combination of the stimuli's expression (fearful or neutral), orientation (upright or inverted), and staircase direction (ascending and descending), with two staircases of each type. Ascending staircases started with an exposure duration of 200 ms whereas descending staircases started at 6826 ms.

Participants were told that each trial would contain a face. They were only instructed to report the location of the face on the screen (left or right). They were not given details about the staircase procedure.

RESULTS

The final five reversal values (durations at which the staircase changed direction) were averaged for each individual staircase. Then, the means were collapsed into four categories: fearful upright, fearful inverted, neutral upright, and neutral inverted faces. Finally, these estimated threshold means were entered into a 2 (expression: fearful, happy) × 2 (orientation: upright, inverted) repeated-measures ANOVA.

Detection thresholds did not differ between conditions: We found no effects of either expression ($F_{(1, 32)} = 0.309, p = .582, \eta p^2 = .010$) or orientation ($F_{(1, 32)} = 0.153, p = .698, \eta p^2 = .005$; Figure 7), nor a significant interaction ($F_{(1, 32)} = 0.126, p = .725, \eta p^2 = .004$). To test whether the evidence supports these null effects, we estimated Bayes factors. They indicated moderate evidence for the null effects of expression ($BF_{01} = 5.403$), orientation ($BF_{01} = 4.544$), and their interaction ($BF_{01} = 3.979$). These results indicate that neither expression nor orientation had an effect on detection threshold estimates.

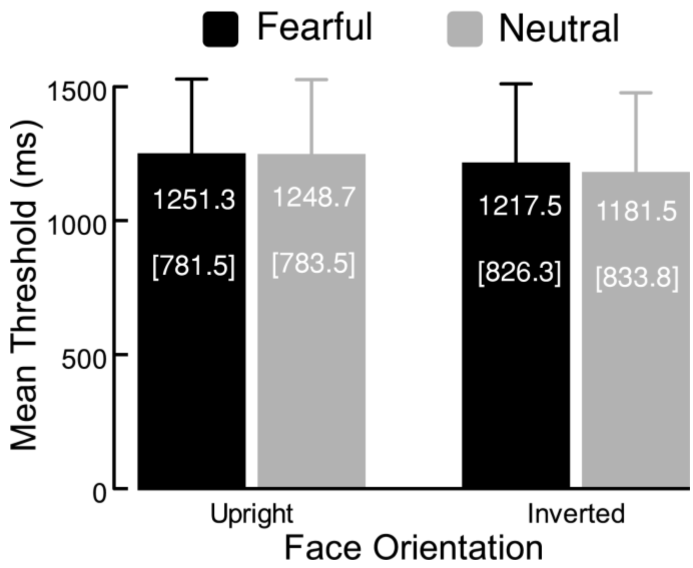

Figure 7. Mean location detection threshold estimates in ms. Each bar shows the mean [SD]. Error bars represent 95% confidence intervals.

DISCUSSION

This experiment addressed the possibility that the pre-determined exposure durations used in our previous experiments missed a crucial range, explaining the absence of facial expression effects. We therefore assessed individual detection thresholds, but still found no differences between expressions, nor any effect of orientation.

Importantly, the mean thresholds found ranged, across conditions, between 1181-1251 ms. This is close to one of the mid-range durations we used in Experiments 1-3 (1350 ms), suggesting that the absence of emotional expression effects in those experiments is unlikely to be due to such effects arising at a mid-point between durations in our set, and confirming that it cannot be attributed to such effects occurring outside the range of durations we used.

One possible interpretation of our findings, in this experiment as well as the preceding ones, is that on those trials where participants were able to report the presence of a face on the correct side, their response was based solely on differences in contrast between screen sides, meaning the present experiment's results correspond to low-level visual thresholds (e.g., to faces' contour or contrast) rather than to high-level features (e.g., emotional expression). However, this would still imply that expression, in itself, does not confer an advantage in making such perceptual decisions. It could also be the case that equating the stimuli in luminance, as we did, substantially decreased low-level visual differences that were driving effects observed in bCFS studies. As described in the Introduction, it has been shown that low-level visual differences may explain shorter breakthrough times of fearful expressions over others (Gray et al., 2013; Hedger et al., 2015; Stein & Sterzer, 2012).

Finally, it is worth noting that the mean threshold durations obtained in this experiment are similar to those obtained by Stein, Hebart, et al. (2011), whose staircase-measures of thresholds for upright vs inverted faces ranged between 1200-1450ms; we did not find a similar effect of orientation, however, which is surprising given that our procedure was stricter, with a higher number of trials and staircases. Interestingly, bCFS studies have typically yielded longer RTs: for example, Jiang et al. (2007) found breakthrough RTs ranging between 1360-1760ms for upright and inverted faces; E. Yang et al. (2007) found breakthrough RTs ranging between ~2000-4000ms when manipulating emotional expression and orientation), which may be explained by the fact that bCFS response times are increased by the need to plan and execute a motor response; this is not the case in non-speeded tasks such as ours.

Consistency of the face-inversion effect (FIE)

In Experiment 2A, we found a FIE for both detection and identification; in Experiment 2B, the FIE was replicated for identification, and was marginal for detection. In contrast, we did not find a detection FIE in Experiments 3 and 4 (which did not include an identification task). It is possible that detection FIE is not very consistent across participants; this may explain why it is not found consistently across experiments. To assess the consistency of the FIE across participants, we estimated the area under the curve (AUC) of each detection (location d') and identification (identification d') measure in the experiments that involved orientation manipulation (Supplementary Material 3). We subtracted AUC values of inverted faces from AUC values of upright faces (for Experiment 4, AUC values were subtracted the other way around). We then counted the frequency (number of participants) in different ranges of differences. As Figure 8 shows, AUC values for the FIE were consistent with the experimental findings reported above: more than half of the participants showed a FIE in Experiments 2A and 2B (for both location d′ and identification d′) whereas in Experiments 3 and 4 most participants did not exhibit a detection FIE. These data suggest that the FIE may not be as robust as reported in past studies. We return to this in the General Discussion.

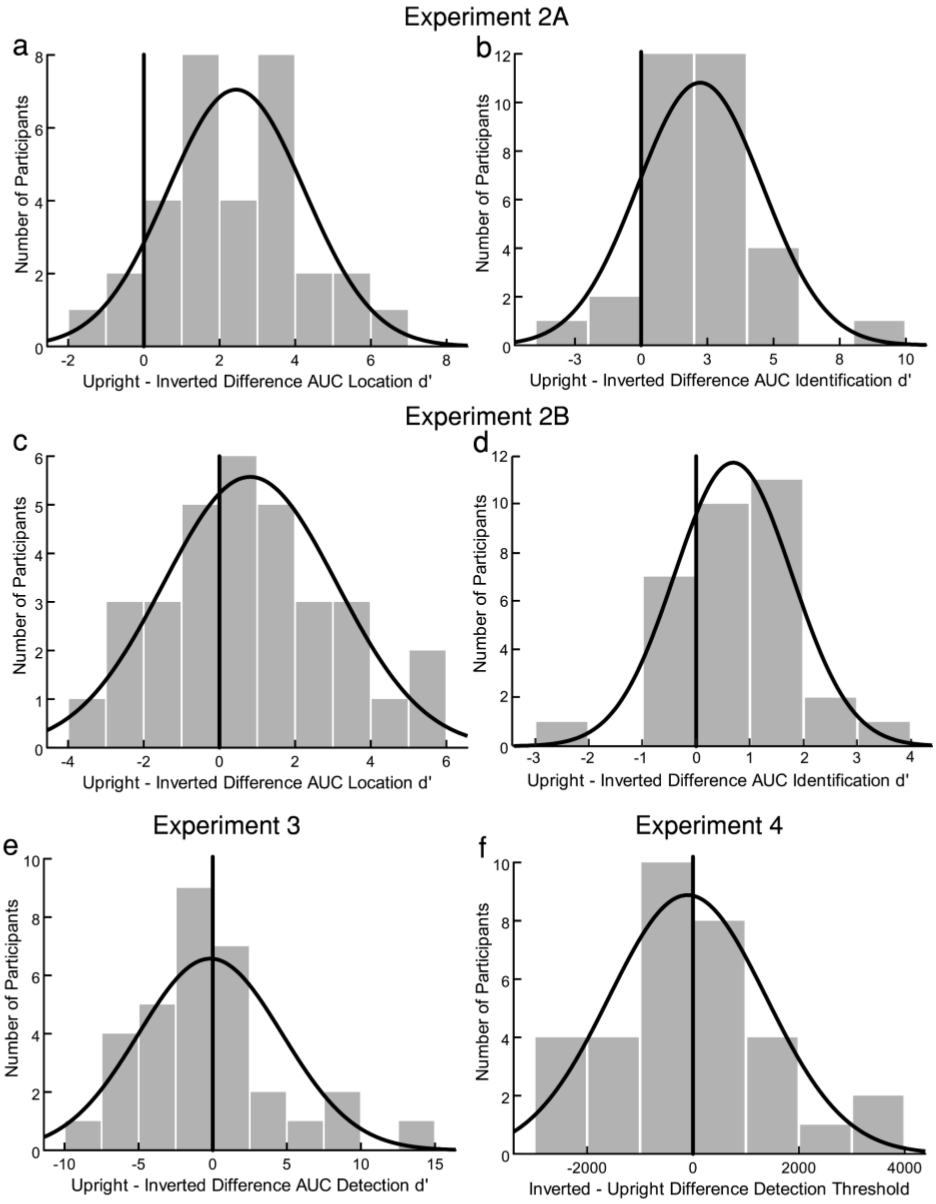

Figure 8. FIE AUC differences across participants. Experiment 2A: (a) Location d′; (b) Expression identification d′. Experiment 2B: (c) Location d′; (d) Expression identification d′. Experiment 3: (e) Location d′. Experiment 4: (f) Detection threshold estimates.

GENERAL DISCUSSION

In the six experiments reported above, we addressed the question of whether emotional facial expressions have an advantage – specifically, a detection sensitivity advantage – over neutral expressions in overcoming interocular suppression. This investigation was motivated by previous work using the bCFS paradigm, which found faster RTs for reporting that emotional (compared to neutral) faces had broken through CFS (Capitão et al., 2014; Hedger et al., 2015; Oliver et al., 2015; Sterzer et al., 2011; E. Yang et al., 2007; Y.-H. Yang & Yeh, 2018; Zhan et al., 2015). As we detailed in the Introduction, such RT results are hard to interpret, because observers can decide how much information to receive before committing to a report; their responses may therefore reflect the emergence of perceptual sensitivity to the suppressed stimulus, but may also be influenced by differences in decision criteria, stimulus identification, and response production processes.

Here, we employed a procedure in which stimuli were presented for predetermined durations and participants provided non-speeded responses. We used a variety of psychophysical approaches – forced choice localisation (Experiments 1A-2B), presence/absence detection (Experiment 3), and staircase-based threshold measurement (Experiment 4) – to directly measure sensitivity for face detection, as well as for identification of emotional expressions' presence. Overall, we found that emotional expressions did not alter detection sensitivity to faces as they break through CFS, which was supported by Bayes Factors analysis (Table 1). Emotional expressions only affected detection sensitivity once (Experiment 2B), and the effect was both very small and contradicted by Bayes Factors, leading us to conclude it was likely a false positive. Emotional expressions did, however, affect identification sensitivity (Table 2): happy expressions were better discriminated from their neutral counterparts than angry (Experiment 1A) and fearful ones (Experiment 1B); the direction of these effects, as noted in the Discussion of Experiment 1A, may reflect characteristics of our stimulus set, suggesting the possibility of similar influences in past bCFS findings and supporting the idea that identification processes may contribute to effects in RT-based paradigms.

Table 1. Summary of main effects of exposure duration, facial expression, and orientation on location sensitivity across experiments.

	Type of effect								
	Effect of exposure			Effect of expression			Effect of orientation		
Exp.	*F*	p	ηp^2	*F*	p	ηp^2	*F*	p	ηp^2
1A	104.5	**< .001***	.771	2.42	.105	.072	-	-	-
1B	111.9	**< .001***	.783	0.103	.887	.003	-	-	-
2A	185.3	**< .001***	.857	2.03	.164	.061	50.4	**< .001***	.62
2B	161	**< .001***	.839	9.406	**.004***	.233	3.511	**.07ᶲ**	.102
3	141.8	**< .001***	.821	0.560	.460	.018	2.191	.149	.066
4	-	-	-	0.309	.582	.001	0.153	.698	.005

Note. Summary of main effects, p-values, and partial eta-squared effect sizes. The single significant effect of expression indicates an advantage of happy over fearful faces. Significant and marginal effects of orientation indicated an advantage of upright over inverted faces. Experiments 1A and 1B did not contain an orientation manipulation; Experiments 3 and 4 did not contain an identification task. Significant and marginal effects are shown in bold. *Significant effects ($p < .05$). ᶲMarginal effect ($p < .1$).

Table 2. Summary of main effects of exposure duration, facial expression, and orientation on expression identification sensitivity across experiments.

	Type of effect								
	Effect of exposure			Effect of expression			Effect of orientation		
Exp.	*F*	p	ηp^2	*F*	p	ηp^2	*F*	p	ηp^2
1A	81.2	**< .001***	.724	4.47	**.043***	.126	-	-	-
1B	76.9	**< .001***	.713	9.86	**.004***	.241	-	-	-
2A	88.5	**< .001***	.74	-	-	-	28.4	**< .001***	.478
2B	176.5	**< .001***	.851	-	-	-	12.8	**< .001***	.292

Note. Summary of main effects, p-values, and partial eta-squared effect sizes. The significant effects of expression indicate an advantage of happy over angry faces in Exp. 1A, and of happy over fearful faces in Exp. 1B. Both significant effects of orientation indicated an advantage of upright over inverted faces. Significant effects are shown in bold. Significant effects ($p < .05$).

As with any null result, it remains possible that our paradigm was not sufficiently sensitive to detect effects of emotional expression on detection sensitivity. We think this is unlikely, though, for several reasons: first, our paradigm is extremely sensitive to small differences in exposure duration. In all our experiments, location d' scores were significantly higher when stimuli were presented for 900 ms compared to 600 ms (all p-values < 0.001; Supplementary Material 4). This 300 ms difference is half the magnitude of the emotional expression effect originally reported by E. Yang et al (2007), where RTs for fearful faces were about 600 ms shorter than for Happy faces. If one stimulus category breaks through CFS 600 ms faster than another, then a paradigm that reliably detects differential perceptual sensitivity to displays whose durations are 300 ms apart would capture such an effect.

Second, our paradigm has been able to confirm differential detection sensitivity for a different effect from the bCFS literature: in a recent study (Lanfranco, Stein, et al., 2022) we used it and found that detection sensitivity is reliably greater for faces making eye contact than for averted-gaze faces, consistent with a different much-replicated bCFS effect (e.g., Akechi et al., 2014; Chen & Yeh, 2012; Stein, Senju, et al., 2011; Yokoyama et al., 2013). Notably, this gaze effect has a similar magnitude in the bCFS literature to that of the emotional expression effect (an RT difference of around 500 to 900 ms for reporting direct- vs averted-gaze faces). To ascertain that the effect of gaze differs significantly from the effect of emotion, we compared them directly by entering the detection sensitivity (location d') data from the current report's Experiment 2A and Experiment 2 in Lanfranco, Stein, et al., 2022 (which had a similar structure) into a mixed ANOVA with 'experiment' as a between-subjects factor (for full details, see Supplementary Material 5). Crucially, we found a significant interaction between experiment and the effect of the manipulation (direct vs averted gaze in the previous study, fearful vs neutral expression in the present one), demonstrating that with very similar setups, our method detects an effect of one manipulation but not the other, supporting the method's ability to distinguish the presence and absence of effects on d'.

Could our paradigm be sufficiently sensitive to detect the effects of gaze and exposure duration on detection sensitivity, but too noisy (and thus not sensitive enough) to detect an effect of emotional expressions on the same measure? There is no way to definitively rule this possibility out, but we believe it is unlikely: As noted above, our experiments employed several psychophysical approaches that (unlike bCFS) are specifically designed to assess detection sensitivity; and although all empirical data contain some noise, our results strongly suggest that the observed sensitivity data have low levels of noise: As demonstrated by all the figures depicting sensitivity data (Figures 2a, 2c, 3a, 3c, 4a, 4c, 5a, 5c, and 6a), d' increases monotonically with display duration; there is not a single case of where an increase in duration is accompanied by a decrease in d' (the kind of up-down reversal that would be expected with noisy data). *Across* experiments, there is one experiment out of six (Experiment 2B) that shows an effect of emotion on detection sensitivity. The presence of cross-experiment noise can cause either of the two types of error, and in light of the low levels of noise within each experiment, we think it is more straightforward to interpret this anomalous result as indicating a single type-1 error (a false positive), rather than a one real effect among five type-2 errors (failures to detect a real effect). The possibility that our paradigm is specifically too noisy to detect the effects of emotion is thus empirically indistinguishable from our main conclusion that in the context of overcoming CFS, emotional facial expression does not affect detection sensitivity.

Finally, some of the present experiments, as well as the aforementioned gaze-direction study using the same paradigm (Lanfranco, Stein, et al., 2022) contained a face-orientation manipulation intended to capture another effect from the bCFS literature – the FIE. Here, the emerging picture broadly supports the paradigm's sensitivity to differential effects, though it is more nuanced: the FIE was consistently observed for identification sensitivity (Table 2) – it replicated across Experiments 2A and 2B (the two experiments that included both an identification task and an orientation manipulation), and was also found in our previous gaze-direction study. This is suggestive evidence for one of our critiques of prior bCFS work: that certain bCFS effects – such as the FIE – may arise, at least in part, thanks to effects on identification rather than detection processes. The FIE was less reliable, however, for location sensitivity: it was significant in Experiment 2A (as well as in our previous gaze-direction study), marginal in Experiment 2B, and absent in Experiments 3 and 4. Overall, therefore, the FIE was seen in the majority of measures meant to capture it (both in the present experiments and when including our other study that used the same paradigm), although it was less consistent than we had believed it to be when we decided to include it in our experimental designs.

This inconsistency, however, is in line with other recent findings: Heyman et al. (2019) also used the FIE to ascertain their paradigm's sensitivity, in a pre-registered report replicating a claim about bCFS and sound symbolism. Notably, they found that obtaining the FIE required 110 participants. As Heyman et al. (2019) note, this is particularly striking because, in a different study, they had obtained strong evidence for the FIE with only 8 participants (see also Moors & Heyman, 2019). Like them, we still think the FIE is likely to be a real effect, but one that is much more heterogenous (or less stable) than previously thought. The findings of Heyman et al. (2019), as well as our current results, thus suggest that the published record's current representation of the FIE's robustness may be exaggerated, possibly reflecting a publication bias (in this context, it is noteworthy that Heyman et al.'s study was a registered report, accepted for publication before the results were known – meaning it was not judged by its results, but solely on the soundness of its methodology).

How can our findings be reconciled with the multiple bCFS reports that emotional expressions overcome suppression faster? As noted, the dependent measure in bCFS studies is RTs; although it is intuitive to interpret faster responses in bCFS as indicating faster emergence of sensitivity to stimuli – especially when the bCFS task included a localisation element – RTs are not a direct measure of sensitivity (unlike the measures used in our present experiments): they are an overall measure arising from all the processes that go into producing a speeded response. A non-comprehensive list of such processes includes, in addition to sensitivity, possible differences in criteria for detection; identification processes and their accompanying biases; and possible effects of stimuli on response production processes. Our experiments were designed to specifically assess the effect of emotional expressions on sensitivity and account for the effects on response bias (in location detection) and criterion (in identification), and our results, therefore, do not refute the findings of bCFS studies; rather, they place important constraints on their interpretation: they indicate that differential responses to emotional expressions in bCFS are driven by some combination of the many other factors that may affect RTs, but not by differential detection sensitivity.

This provides a starting point for further studies aimed at identifying the specific process – or processes – that is responsible for these bCFS findings. We note that the specific combination of processes which drives emotion effects in bCFS may differ from the combinations that underlie other reported bCFS effects. Different stimulus manipulations may affect different processes, precluding a single, unified account for all the effects observed in bCFS studies.

The data regarding different stimulus manipulations in our psychophysical studies to date support this idea: as noted above, we have found that gaze direction affects both location detection and identification sensitivity under CFS (Lanfranco, Stein, et al., 2022), paralleling the finding obtained in bCFS studies; in contrast, the FIE is consistently present for identification sensitivity (in both Lanfranco, Stein, et al., 2022 and the present study), but inconsistent for detection sensitivity; and finally, emotional expression, consistently, does not have an effect on detection sensitivity, but can affect expression identification that the presence of a FIE at individual level does not. We ran additional analyses to examine whether an FIE on detection sensitivity might be a prerequisite for a detection advantage for fearful faces, but found no evidence to support this possibility (see Supplementary Material 6). These different findings suggest that different stimulus manipulations have distinct effects on different processes involved in overcoming CFS; although they may all lead to RT effects in bCFS, the RT measure is unable to distinguish the different underlying mechanisms of each effect. Preferential breakthrough as indicated by bCFS effects therefore points toward the need to pinpoint the underlying processes. Some of these processes – those related to perceptual sensitivity and decisional factors – can be isolated with the psychophysical methods we employed here. Others (e.g., those related to response production) may require modelling of response production in RT data or the use of physiological measures.

Another way in which our experiments differ from bCFS studies is that we used objective measures of perceptual performance, whereas in bCFS participants are instructed to respond as soon as they become (subjectively) aware of the stimulus. Could this distinction account for the difference between the present and previous findings? We do not think so: objective performance and subjective awareness are indeed distinct and can sometimes be dissociated – above-chance performance can occur without subjective awareness, as is, for example, the case in blindsight (e.g., Kentridge et al., 1999). This, however, does not imply the opposite dissociation: outside of perceptual illusions – and certainly in the current context of overcoming suppression – it is hard to conceptualise the presence of subjective perceptual awareness that confers no ability to make veridical, objectively correct perceptual judgments. Indeed, bCFS studies often use an (objective) 2AFC or 4AFC spatial localisation task to verify that the observers' stimulus-awareness reports are veridical (E. Yang et al., 2007; Capitão et al., 2014; Yang & Yeh, 2018). This ability to make forced-choice judgments – i.e., to have objective perceptual sensitivity – is assumed to arise from the stimulus reaching awareness; there would be no point in using it in the context of bCFS if it were an indication, for example, of unconscious perception. Stimulus categories that yield faster RTs in bCFS are thus thought to enjoy faster access to awareness, enabling such sensitivity. If this is true, then those stimulus categories should also yield better perceptual sensitivity when this is measured in a non-speeded task, independently of the many other factors that may influences RTs.

Furthermore, despite their instructions, bCFS tasks do not provide an isolated measure of subjective awareness. Considering RTs in this paradigm as reflecting either objective or subjective responses is an oversimplification: as described in detail in the Introduction, it is possible that two stimuli would reach awareness with the same latency, and nonetheless yield different RTs due to differential influences on the many additional processes that are involved in response generation. One cannot assume that participants are aware of these influences, or that their response latencies purely reflect the time a stimulus reached awareness just because of the instructions' phrasing.

It is, however, possible to assess subjective perceptual experience using various measures such as confidence ratings or perceptual awareness scale (PAS) reports, with discrete, continuous, or even combined discrete-continuous scales (Sandberg et al., 2010; Wierzchoń et al., 2019); it is also possible to assess the correspondence between such subjective reports and objective performance to obtain measures of metacognitive sensitivity and efficiency (Fleming, 2017; Fleming & Lau, 2014). If we had found a reliable difference in objective detection sensitivity between expressions, it would have been fruitful to follow this up by assessing the correspondence between psychophysical detection sensitivity and subjective awareness, using such measures. This would involve collecting, on each trial, a measure of subjective awareness, such as a PAS report, in addition to the objective performance report (e.g., Lanfranco et al., 2021). We note that this is not straightforward from a technical perspective: the large number of trials required for psychophysical assessment, and the long duration of individual trials in such an investigation (where adding a subjective report would make each trial even longer), necessitate that one first ascertain the exposure durations at which objective performance differs for different expressions; using a small set of relevant durations would make the addition of subjective experience measures feasible. We did not, however, find the requisite differences: rather, having tried localisation sensitivity, presence/absence sensitivity, and staircase-based thresholds, we eventually concluded there was no objective detection sensitivity difference to be found.

Recent studies have suggested that the advantage of emotional over non-emotional expressions found with the bCFS procedure may have been due to low-level features like spatial frequency and contrast (Gelbard-Sagiv et al., 2016; Schlossmacher et al., 2017; Stein et al., 2018; for a review, see Lanfranco et al., 2023). It could be argued that the fact that we only controlled luminance and still did not find an advantage of emotional expressions suggests that low-level differences are not the only cause of previously reported effects. In the ongoing debate on the role of low-level features in the effects of emotional expressions, it is worth noting recent studies, which have shown that contrast normalisation of face stimuli may, in fact, artifactually increase the detectability of fearful expressions (Webb et al., 2020; Webb & Hibbard, 2020). These findings raise additional concerns regarding which low-level features should be equated in face detection studies. Regardless of the role and magnitude of low-level feature differences, such differences may have been reduced in our studies, as we employed a substantially larger number of face stimuli than in past bCFS studies (minimising the idiosyncratic influence of any specific stimulus), and selected them based on published norms to match emotion intensity and minimise identifiability differences between conditions.

It is worth noting that our experiments assessed sensitivity to the presence of an emotional expression, but not the ability to correctly classify which emotional expression was shown (except in Experiment 2B, where participants categorised faces as happy or fearful; it could, however, be argued that this only requires the ability to reliably identify one of the expressions). Although this design was appropriate for addressing our main aim (to examine whether the very presence of certain expressions affects suppression), we note that emotion detection and categorisation are dissociable processes (Sweeny et al., 2013). It would therefore be interesting to extend the present findings – especially those concerning identification sensitivity – by using a task in which different emotional expressions are used within the same blocks, and observers are explicitly required to label the emotions.

In closing, we reiterate that while our findings rule out differential detection sensitivity as a basis for faster breakthrough of emotional expressions, they do not rule out the faster responses themselves (which may be due to several other factors). We are also not suggesting that processing of emotional expressions can only happen consciously. Indeed, there is a great deal of neural and physiological evidence for the processing of emotional facial expressions under interocular (Pasley et al., 2004; Williams et al., 2004) and other methods of suppression (Dimberg et al., 2000; Esteves et al., 1994; Morris et al., 1998), as well as evidence for unconscious processing of emotional information more generally (Homan et al., 2021; Nasrallah et al., 2009; Raio et al., 2012; Tooley et al., 2017). Our findings bear on the further idea that such unconscious processing promotes faster access to awareness, and highlight the need for evidence that would support it using stringent psychophysical procedures and measures.

Acknowledgements

This work was supported by five University of Edinburgh PPLS Research Support Grants, an ANID/CONICYT PhD studentship awarded at the University of Edinburgh, and a postdoctoral fellowship awarded at Karolinska Institutet to R.C.L., and by an ESRC Future Research Leaders award (no. ES/L01064X/1) to H.R.; D.C. was supported by a European Research Council Grant (XSPECT - DLV-692739). The authors thank Alisdair Tullo for his technical advice and help.

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. https://doi.org/10.1016/j.cognition.2015.06.016

Akechi, H., Stein, T., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2014). Absence of Preferential Unconscious Processing of Eye Contact in Adolescents With Autism Spectrum Disorder. *Autism Research*, *7*(5), 590–597. https://doi.org/10.1002/aur.1397

Almeida, J., Pajtas, P. E., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2013). Affect of the unconscious: Visually suppressed angry faces modulate our decisions. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(1), 94–101. https://doi.org/10.3758/s13415-012-0133-7

Alpers, G. W., & Gerdes, A. B. M. (2007). Here is looking at you: Emotional faces predominate in binocular rivalry. *Emotion*, *7*(3), 495–506. https://doi.org/10.1037/1528-3542.7.3.495

Balconi, M., & Pozzoli, U. (2003). Face-selective processing and the effect of pleasant and unpleasant emotional expressions on ERP correlates. *International Journal of Psychophysiology*, *49*(1), 67–74.

https://doi.org/10.1016/S0167-8760(03)00081-3

Biderman, N., & Mudrik, L. (2018). Evidence for Implicit—But Not Unconscious—Processing of Object-Scene Relations. *Psychological Science*, *29*(2), 266–277. https://doi.org/10.1177/0956797617735745

Capitão, L. P., Underdown, S. J. V., Vile, S., Yang, E., Harmer, C. J., & Murphy, S. E. (2014). Anxiety increases breakthrough of threat stimuli in continuous flash suppression. *Emotion*, *14*(6), 1027–1036. https://doi.org/10.1037/a0037801

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

Carmel, D., Arcaro, M., Kastner, S., & Hasson, U. (2010). How to Create and Use Binocular Rivalry. *Journal of Visualized Experiments*, *45*, Article e2030. https://doi.org/10.3791/2030

Chen, Y.-C., & Yeh, S.-L. (2012). Look into my eyes and I will see you: Unconscious processing of human gaze. *Consciousness and Cognition*, *21*(4), 1703–1710. https://doi.org/10.1016/j.concog.2012.10.001

Costello, P., Jiang, Y., Baartman, B., McGlennen, K., & He, S. (2009). Semantic and subword priming during binocular suppression. *Consciousness and Cognition*, *18*(2), 375–382. https://doi.org/10.1016/j.concog.2009.02.003

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious Facial Reactions to Emotional Facial Expressions. *Psychological Science*, *11*(1), 86-89. https://doi.org/10.1111/1467-9280.00221

Esteves, F., Parra, C., Dimberg, U., & Öhman, A. (1994). Nonconscious associative learning: Pavlovian conditioning of skin conductance responses to masked fear-relevant facial stimuli. *Psychophysiology*, *31*(4), 375–385. https://doi.org/10.1111/j.1469-8986.1994.tb02446.x

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is 'special' about face perception? *Psychological Review*, *105*(3), 482–498. https://doi.org/10.1037/0033-295X.105.3.482

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fleming, S. M. (2017). HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, *2017*(1), Article nix007. https://doi.org/10.1093/nc/nix007

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8, Article 443. https://doi.org/10.3389/fnhum.2014.00443*.

Fox, E. (2002). Processing emotional facial expressions: The role of anxiety and awareness. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(1), 52–63. https://doi.org/10.3758/CABN.2.1.52

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*(1), 61–92. https://doi.org/10.1080/026999300378996

Fredrikson, M., Furmark, T., Olsson, M. T., Fischer, H., Andersson, J., & Långström, B. (1998). Functional neuroanatomical correlates of electrodermal activity: A positron emission tomographic study. *Psychophysiology*, *35*(2), 179–185. https://doi.org/10.1111/1469-8986.3520179

Gayet, S., Paffen, C. L. E., Belopolsky, A. V., Theeuwes, J., & Van der Stigchel, S. (2016). Visual input signaling threat gains preferential access to awareness in a breaking continuous flash suppression paradigm. *Cognition*, *149*, 77–83. https://doi.org/10.1016/j.cognition.2016.01.009

Gayet, S., & Stein, T. (2017). Between-Subject Variability in the Breaking Continuous Flash Suppression Paradigm: Potential Causes, Consequences, and

Solutions. *Frontiers in Psychology*, *8, Article 437*. https://doi.org/10.3389/fpsyg.2017.00437
Gayet, S., Van der Stigchel, S., & Paffen, C. L. E. (2014). Breaking continuous flash suppression: Competing for consciousness on the pre-semantic battlefield. *Frontiers in Psychology*, *5, Article 460*. https://doi.org/10.3389/fpsyg.2014.00460
Gelbard-Sagiv, H., Faivre, N., Mudrik, L., & Koch, C. (2016). Low-level awareness accompanies "unconscious" high-level processing during continuous flash suppression. *Journal of Vision*, *16*(1), Article 3. https://doi.org/10.1167/16.1.3
Gelman, A., & Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642
Goeleven, E., Raedt, R. D., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition and Emotion*, *22*(6), 1094–1118. https://doi.org/10.1080/02699930701626582
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*, *13*(3), 537–544. https://doi.org/10.1037/a0031403
Grill-Spector, K., Weiner, K. S., Kay, K., & Gomez, J. (2017). The Functional Neuroanatomy of Human Face Perception. *Annual Review of Vision Science*, *3*(1), 167–196. https://doi.org/10.1146/annurev-vision-102016-061214
Hassin, R. R. (2013). Yes It Can: On the Functional Abilities of the Human Unconscious. *Perspectives on Psychological Science*, *8*(2), 195–207. https://doi.org/10.1177/1745691612460684
Hedger, N., Adams, W. J., & Garner, M. (2015). Fearful faces have a sensory advantage in the competition for awareness. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1748–1757. https://doi.org/10.1037/xhp0000127
Hedger, N., Gray, K. L. H., 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*, *142*(9), 934–968. https://doi.org/10.1037/bul0000054
Hesselmann, G., & Moors, P. (2015). Definitely maybe: Can unconscious processes perform the same functions as conscious processes? *Frontiers in Psychology*, *6, Article 584*. https://doi.org/10.3389/fpsyg.2015.00584
Heyman, T., Maerten, A.-S., Vankrunkelsven, H., Voorspoels, W., & Moors, P. (2019). Sound-Symbolism Effects in the Absence of Awareness: A Replication Study. *Psychological Science*, *30*(11), 1638–1647. https://doi.org/10.1177/0956797619875482
Homan, P., Lau, H. L., Levy, I., Raio, C. M., Bach, D. R., Carmel, D., & Schiller, D. (2021). Evidence for a minimal role of stimulus awareness in reversal of threat learning. *Learning & Memory*, *28*(3), 95–103. https://doi.org/10.1101/lm.050997.119
Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052
JASP Team. (2020). *JASP* (0.12.2) [Computer software]. https://jasp-stats.org/
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. https://doi.org/10.1111/j.1467-9280.2007.01902.x
Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology*, *38*(2), 175-178.

https://doi.org/10.1111/1469-8986.3820175
Jusyte, A., Mayer, S. V., Künzel, E., Hautzinger, M., & Schönenberg, M. (2015). Unemotional traits predict early processing deficit for fearful expressions in young violent offenders: An investigation using continuous flash suppression. *Psychological Medicine*, *45*(2), 285–297. https://doi.org/10.1017/S0033291714001287
Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (1999). Attention without awareness in blindsight. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1430), 1805–1811. https://doi.org/10.1098/rspb.1999.0850
Kobylka, F., Persike, M., & Meinhardt, G. (2017). Object localization does not imply awareness of object category at the break of continuous flash suppression. *Frontiers in Human Neuroscience*, *11, Article 312*. *https://doi.org/10.3389/fnhum.2017.00312*
Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394–421. https://doi.org/10.1016/j.biopsycho.2010.03.010
Krysko, K. M., & Rutherford, M. D. (2009). The face in the crowd effect: Threat-detection advantage with perceptually intermediate distractors. *Visual Cognition*, *17*(8), 1205–1217. https://doi.org/10.1080/13506280902767789
Lanfranco, R. C. (2021). *Assessing the influence of emotional expressions on perceptual sensitivity to faces overcoming interocular suppression*. OSF. https://osf.io/m83qv/
Lanfranco, R. C., Canales-Johnson, A., Rabagliati, H., Cleeremans, A., & Carmel, D. (2021). *Minimal exposures reveal visual processing priorities*. bioRxiv. https://doi.org/10.1101/2021.10.06.462913
Lanfranco, R. C., Rabagliati, H., & Carmel, D. (2023). The importance of awareness in face processing: A critical review of interocular suppression studies. *Behavioural Brain Research*, *437*, Article 114116. https://doi.org/10.1016/j.bbr.2022.114116
Lanfranco, R. C., Stein, T., Rabagliati, H., & Carmel, D. (2022). Gaze direction and face orientation modulate perceptual sensitivity to faces under interocular suppression. *Scientific Reports*, *12*(1), Article 7640. https://doi.org/10.1038/s41598-022-11717-4
Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x
Li, Y., & Li, S. (2015). Contour integration, attentional cuing, and conscious awareness: An investigation on the processing of collinear and orthogonal contours. *Journal of Vision*, *15*(16), Article 10. https://doi.org/10.1167/15.16.10
Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and 'nonparametric' indexes. *Psychological Bulletin*, *107*(3), 401–413. https://doi.org/10.1037/0033-2909.107.3.401
Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge University Press.
Madipakkam, A. R., Rothkirch, M., Guggenmos, M., Heinz, A., & Sterzer, P. (2015). Gaze Direction Modulates the Relation between Neural Responses to Faces and Visual Awareness. *Journal of Neuroscience*, *35*(39), 13287–13299. https://doi.org/10.1523/JNEUROSCI.0815-15.2015
Mihalache, D., Lamer, S. A., Allen, J., Maher, M., & Sweeny, T. D. (2021). Anger bias in the evaluation of crowds. *Journal of Experimental Psychology: General, 150*(9), 1870-1889. https://doi.org/10.1037/xge0001025
Moors, P., Boelens, D., van Overwalle, J., & Wagemans, J. (2016). Scene Integration Without Awareness: No Conclusive Evidence for Processing Scene Congruency During

Continuous Flash Suppression. *Psychological Science*, *27*(7), 945–956. https://doi.org/10.1177/0956797616642525
Moors, P., & Hesselmann, G. (2018). A critical reexamination of doing arithmetic nonconsciously. *Psychonomic Bulletin & Review*, *25*(1), 472–481. https://doi.org/10.3758/s13423-017-1292-x
Moors, P., & Heyman, T. (2019). Sequential testing, replication, and the unconscious [Blog]. *The 100% CI*. http://www.the100.ci/2019/11/06/sequential-testing-replication-and-the-unconscious/
Moors, P., Wagemans, J., & de-Wit, L. (2016). Faces in commonly experienced configurations enter awareness faster due to their curvature relative to fixation. *PeerJ*, *4*, e1565. https://doi.org/10.7717/peerj.1565
Moors, P., Wagemans, J., van Ee, R., & de-Wit, L. (2016). No evidence for surface organization in Kanizsa configurations during continuous flash suppression. *Attention, Perception, & Psychophysics*, *78*(3), 902–914. https://doi.org/10.3758/s13414-015-1043-x
Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*(6684), Article 6684. https://doi.org/10.1038/30976
Mudrik, L., Breska, A., Lamy, D., & Deouell, L. Y. (2011). Integration Without Awareness: Expanding the Limits of Unconscious Processing. *Psychological Science*, *22*(6), 764–770. https://doi.org/10.1177/0956797611408736
Münkler, P., Rothkirch, M., Dalati, Y., Schmack, K., & Sterzer, P. (2015). Biased Recognition of Facial Affect in Patients with Major Depressive Disorder Reflects Clinical State. *PLOS ONE*, *10*(6), e0129863. https://doi.org/10.1371/journal.pone.0129863
Nasrallah, M., Carmel, D., & Lavie, N. (2009). Murder, she wrote: Enhanced sensitivity to negative word valence. *Emotion*, *9*(5), 609–618. https://doi.org/10.1037/a0016305
Oliver, L. D., Mao, A., & Mitchell, D. G. V. (2015). "Blindsight" and subjective awareness of fearful faces: Inversion reverses the deficits in fear perception associated with core psychopathic traits. *Cognition and Emotion*, *29*(7), 1256–1277. https://doi.org/10.1080/02699931.2014.976182
Oruc, I., Balas, B., & Landy, M. S. (2019). Face perception: A brief journey through recent discoveries and current directions. *Vision Research*, *157*, 1–9. https://doi.org/10.1016/j.visres.2019.06.005
Paffen, C. L. E., Gayet, S., Heilbron, M., & Stigchel, S. V. der. (2018). Attention-based perceptual learning does not affect access to awareness. *Journal of Vision*, *18*(3), Article 7. https://doi.org/10.1167/18.3.7
Pasley, B. N., Mayes, L. C., & Schultz, R. T. (2004). Subcortical Discrimination of Unperceived Objects during Binocular Rivalry. *Neuron*, *42*(1), 163-172. https://doi.org/10.1016/S0896-6273(04)00155-2
Peters, M. A. K., Ro, T., & Lau, H. (2016). Who's afraid of response bias? *Neuroscience of Consciousness*, *2016*(1), Article niw001. https://doi.org/10.1093/nc/niw001
Pournaghdali, A., & Schwartz, B. L. (2020). Continuous flash suppression: Known and unknowns. *Psychonomic Bulletin & Review*, *27*(6), 1071–1103. https://doi.org/10.3758/s13423-020-01771-2
Rabagliati, H., Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. *Journal of Experimental Psychology: General*, *147*(2), 190–208. https://doi.org/10.1037/xge0000348
Raio, C. M., Carmel, D., Carrasco, M., & Phelps, E. A. (2012). Nonconscious fear is quickly acquired but swiftly forgotten. *Current Biology*, *22*(12), R477–R479. https://doi.org/10.1016/j.cub.2012.04.023
Salomon, R., Lim, M., Herbelin, B., Hesselmann, G., & Blanke, O. (2013). Posing for awareness: Proprioception modulates access to visual

consciousness in a continuous flash suppression task. *Journal of Vision*, *13*(7), Article 2. https://doi.org/10.1167/13.7.2

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069–1078. https://doi.org/10.1016/j.concog.2009.12.013

Schlossmacher, I., Junghöfer, M., Straube, T., & Bruchmann, M. (2017). No differential effects to facial expressions under continuous flash suppression: An event-related potentials study. *NeuroImage*, *163*, 276–285. https://doi.org/10.1016/j.neuroimage.2017.09.034

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Stein, T. (2019). The Breaking Continuous Flash Suppression Paradigm: Review, evaluation, and outlook. In G. Hesselmann (Ed.), *Transitions between Consciousness and Unconsciousness* (pp. 1–38). Routledge. https://www.taylorfrancis.com/chapters/breaking-continuous-flash-suppression-paradigm-timo-stein/10.4324/9780429469688-1

Stein, T., Awad, D., Gayet, S., & Peelen, M. V. (2018). Unconscious processing of facial dominance: The role of low-level factors in access to awareness. *Journal of Experimental Psychology: General*, *147*(11), e1–e13. https://doi.org/10.1037/xge0000521

Stein, T., Hebart, M. N., & Sterzer, P. (2011). Breaking Continuous Flash Suppression: A New Measure of Unconscious Processing during Interocular Suppression? *Frontiers in Human Neuroscience*, *5*, Article 167. https://doi.org/10.3389/fnhum.2011.00167

Stein, T., & Peelen, M. V. (2021). Dissociating conscious and unconscious influences on visual detection effects. *Nature Human Behaviour*, *5*(5), 612–624. https://doi.org/10.1038/s41562-020-01004-5

Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition*, *119*(2), 307–311. https://doi.org/10.1016/j.cognition.2011.01.008

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

Sterzer, P., Hilgenfeldt, T., Freudenberg, P., Bermpohl, F., & Adli, M. (2011). Access of emotional information to visual awareness in patients with major depressive disorder. *Psychological Medicine*, *41*(8), 1615–1624. https://doi.org/10.1017/S0033291710002540

Sweeny, T. D., Suzuki, S., Grabowecky, M., & Paller, K. A. (2013). Detecting and categorizing fleeting emotions in faces. *Emotion*, *13*(1), 76–91. https://doi.org/10.1037/a0029193

Sylvers, P. D., Brennan, P. A., & Lilienfeld, S. O. (2011). Psychopathic Traits and Preattentive Threat Processing in Children: A Novel Test of the Fearlessness Hypothesis. *Psychological Science*, *22*(10), 1280–1287. https://doi.org/10.1177/0956797611420730

The jamovi project. (2020). *Jamovi* (1.2.17) [Computer software]. Retrieved from https://www.jamovi.org

Tooley, M. D., Carmel, D., Chapman, A., & Grimshaw, G. M. (2017). Dissociating the physiological components of unconscious emotional responses. *Neuroscience of Consciousness*, *2017*(1), Article nix021. https://doi.org/10.1093/nc/nix021

Troiani, V., Price, E. T., & Schultz, R. T. (2014). Unseen fearful faces promote amygdala guidance of attention. *Social Cognitive and Affective Neuroscience*, *9*(2), 133–140. https://doi.org/10.1093/scan/nss116

Webb, A. L. M., & Hibbard, P. B. (2020). Suppression durations for facial expressions under breaking continuous flash suppression: Effects of faces' low-level image properties. *Scientific Reports*, *10*(1), 17427. https://doi.org/10.1038/s41598-020-74369-2

Webb, A. L. M., Hibbard, P. B., & O'Gorman, R. (2020). Contrast normalisation masks natural expression-related differences and artificially enhances the perceived salience of fear expressions. *PLOS ONE*, *15*(6), Article e0234513. https://doi.org/10.1371/journal.pone.0234513

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked Presentations of Emotional Facial Expressions Modulate Amygdala Activity without Explicit Knowledge. *Journal of Neuroscience*, *18*(1), 411–418. https://doi.org/10.1523/JNEUROSCI.18-01-00411.1998

Wickens, T. D. (2001). Elementary Signal Detection Theory. In *Elementary Signal Detection Theory*. Oxford University Press. *https://doi.org/10.1093/acprof:oso/9780195092509.001.0001*

Wierzchoń, M., Anzulewicz, A., Hobot, J., Paulewicz, B., & Sackur, J. (2019). In search of the optimal measure of awareness: Discrete or continuous? *Consciousness and Cognition*, *75*, Article 102798. https://doi.org/10.1016/j.concog.2019.102798

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

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. *Journal of Neuroscience*, *24*(12), 2898–2904. https://doi.org/10.1523/JNEUROSCI.4977-03.2004

Wilson, E., & MacLeod, C. (2003). Contrasting two accounts of anxiety-linked attentional bias: Selective attention to varying levels of stimulus threat intensity. *Journal of Abnormal Psychology*, *112*(2), 212–218. https://doi.org/10.1037/0021-843X.112.2.212

Winer, E. S., & Snodgrass, M. (2015). Signal Detection Theory. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception* (pp. 717-733). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199600472.013.016

Yang, E., Brascamp, J., Kang, M.-S., & Blake, R. (2014). On the use of continuous flash suppression for the study of visual processing outside of awareness. *Frontiers in Psychology*, *5, Article 724*. https://doi.org/10.3389/fpsyg.2014.00724

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, *7*(4), 882–886. https://doi.org/10.1037/1528-3542.7.4.882

Yang, Y.-H., & Yeh, S.-L. (2018). Can emotional content be extracted under interocular suppression? *PLOS ONE*, *13*(11), Article e0206799. https://doi.org/10.1371/journal.pone.0206799

Yokoyama, T., Noguchi, Y., & Kita, S. (2013). Unconscious processing of direct gaze: Evidence from an ERP study. *Neuropsychologia*, *51*(7), 1161-1168. https://doi.org/10.1016/j.neuropsychologia.2013.04.002

Zhan, M., Hortensius, R., & Gelder, B. de. (2015). The Body as a Tool for Anger Awareness—Differential Effects of Angry Facial and Bodily Expressions on Suppression from Awareness. *PLOS ONE*, *10*(10), Article e0139768. https://doi.org/10.1371/journal.pone.0139768

Zhou, G., Zhang, L., Liu, J., Yang, J., & Qu, Z. (2010). Specificity of face processing without awareness. *Consciousness and Cognition*, *19*(1), 408–412. https://doi.org/10.1016/j.concog.2009.12.009