Searching for Emotion: A Top-Down Set Governs Attentional Orienting to Facial Expressions

Hannah L. Delchau, Bruce K. Christensen, Ottmar V. Lipp, Richard O'Kearney, Kavindu H. Bandara, Nicole Tan, Hana Yabuki, & Stephanie C. Goodhew[1]

Research School of Psychology, The Australian National University

School of Psychology, Curtin University

Word count: (main text excluding abstract and references): 10,112

Corresponding Author: Stephanie C. Goodhew

Address: Research School of Psychology (Building 39)

The Australian National University, Canberra, 2601

Email: stephanie.goodhew@anu.edu.au

Abstract

Research indicates that humans orient attention toward facial expressions of emotion. Orienting to facial expressions has typically been conceptualised as due to bottom-up attentional capture. However, this overlooks the contributions of top-down attention and selection history. In the present study, across four experiments, these three attentional processes were differentiated using a variation of the dot-probe task, in which participants were cued to attend to a happy or angry face on each trial. Results show that attention toward facial expressions was not exclusively driven by bottom-up attentional capture; instead, participants could shift their attention toward both happy and angry faces in a top-down manner. This effect was not found when the faces were inverted, indicating that top-down attention relies on holistic processing of the face. In addition, no evidence of selection history was found (i.e., no improvement on repeated trials or blocks of trials in which the task was to orient to the same expression). Altogether, these results suggest that humans can use top-down attentional control to rapidly orient attention to emotional faces.

Keywords: selective attention, spatial attention, dot-probe, facial expressions, emotion.

Searching for Emotion: A Top-Down Set Governs Attentional Orienting to Facial Expressions

You see two people before you; one with an angry expression and one with a happy expression. Which face captures your attention? And why? Despite decades of research on attentional orienting to emotional faces, the answer to these questions remains elusive. This is an important question to address because the ability and tendency to orient to faces of particular expressions have implications for a person's social functioning, and atypical orienting is linked to disorders such as social anxiety (e.g., Grafton & MacLeod, 2016; Lin, Hofmann, Qian, Kind, & Yu, 2016; Pishyar, Harris, & Menzies, 2004). The purpose of the current study, therefore, was to determine the relative contributions of different core mechanisms of attention to emotional face processing. The necessity of visual selection is ever-present as we are constantly bombarded with visual information, only a small amount of which our cognitive systems can process at any given moment (e.g., Broadbent, 1958; Neisser, 1967; Treisman, 1960). Attention plays an important role in selecting particular information for processing at the expense of other information. But what mechanisms guide selective attention? Traditionally, such mechanisms have been categorised as either top-down or bottom-up. Top-down attention refers to the voluntary allocation of attention toward objects, features, or spatial locations (e.g., searching for a friend's blue coat in a crowd) (Corbetta & Shulman, 2002), whereas bottom-up attention refers to the involuntary capture of attention by physically salient stimuli (e.g., a flashing light) (Corbetta & Shulman, 2002). A third mechanism of selective attention has also now been recognised; selection history (Awh, Belopolsky, & Theeuwes, 2012). That is, attention can be captured by a visual stimulus due to its past selection or reward history, even if it no longer matches the individual's top-down goals. Historically, research on attentional orienting to faces displaying emotion has typically assumed, either implicitly or explicitly, that orienting is driven by bottom-up factors (e.g.,

Hansen & Hansen, 1988; Williams, Moss, Bradshaw, & Mattingley, 2005). However, theory

and research from the field of visual attention indicates that the focus of attention at a given

point in time is a product of both top-down and bottom-up processes (e.g., Folk &

Remington, 2010; Wolfe 1994). This challenges the assumption of bottom-up factors driving

attentional orienting to faces as, instead, it seems likely that top-down processes are

implicated. Moreover, more recently, the role of selection history in attentional orienting has

been distinguished from 'pure' top-down processes. Therefore, the present study

systematically examined the relative contribution of top-down attention, bottom-up attention,

and selection history in orienting toward emotional faces.

Bottom-Up Attention

A common conceptualisation of attentional selection is that preferential attention is

directed toward the most salient visual stimuli, locations, or visual features that evoke the

strongest neural response within the visual field (Desimone & Duncan, 1995; Koch &

Ullman, 1985). A widely-held view in the literature is that emotional faces capture attention

in a bottom-up manner as they are more salient than neutral faces. Traditionally, it has been

posited that *threatening* expressions hold special significance for attentional selection

(Hansen & Hansen, 1988; Lipp, Price, & Tellegen, 2009a, 2009b). For instance, in Hansen

and Hansen's (1988) classic face-in-the-crowd experiment, participants' reaction times (RTs)

to detect an angry face did not increase with increasing set size, indicating that angry faces

"popped-out" of the array (however, see Purcell, Stewart, & Skov, 1996, for an alternative

account of this result). This is consistent with an evolutionary account that being able to

respond to threat through bottom-up processing is rapid and automatic and, therefore,

adaptive as it allows individuals to respond quickly and protect themselves (LoBue, Rakison,

& DeLoache, 2010; Mogg & Bradley, 1998; Öhman, 2007). However, other research has

suggested that attention is not specifically oriented towards threat, but towards emotion more

broadly. For example, studies have also shown a *happy* face advantage (Becker, Anderson, Mortensen, Neufeld, & Neel, 2011), though other research has found that happy and angry face advantages in visual search depend on the type of face database employed (Savage, Lipp, Craig, Becker, & Horstmann, 2013). The lack of consistent findings may indicate that attentional capture by emotional faces is associated with low-level perceptual differences across face databases, rather than bottom-up attention to the expressions themselves. Since bottom-up attentional orienting is conceptualised as context-invariant, the diversity of results observed in this literature undermines the notion that facial expressions represent a basic stimulus property that consistently orient attention in a bottom-up way. While there may be some role for bottom-up processes in orienting to faces of particular facial expressions, there is reason to believe that other factors are also at work.

Top-Down Attention

There are issues with bottom-up attentional accounts of orienting to faces of different emotional expressions, as described above. Moreover, a large body of research using basic visual features also indicates that top-down attention has a powerful influence over the selection of stimuli in our visual field, as described in this section. At the very least, this likely implicates a role for the interplay between top-down and bottom-up attention, if not for a dominant role of top-down attention in attentional orienting to faces.

The power of top-down attention is demonstrated by the classic and robust endogenous cueing procedure described by Posner (1980). In this task, on each trial, a cue (e.g., an arrow or digit) is used to inform participants of the likely location of a to-be-presented target. The cue is not intrinsically salient, as is the case for exogenous attentional cueing. Yet participants reliably utilise the informativeness of the cue, as evidenced by the fact that they respond faster and more accurately to the target in the cued location compared

with the un-cued location (Posner, 1980; Theeuwes & Van der Burg, 2007). This is thought to reflect participants using their top-down set to guide attention to a location in space.

In addition to top-down guidance of attention to spatial locations, top-down attention can also be guided to particular object features. According to contingent capture theory, visual cues, that would otherwise be salient in a bottom-up fashion, will only capture spatial attention if they match the top-down attentional set of the participant. In a seminal demonstration of this, Folk, Remington, and Johnston (1992) conducted an experiment in which participants responded to the presence of an "X" or "+" in a stimulus array. These targets could either be defined by onset, whereby a single target would appear on each trial, or by its red colour amongst white distractors. On some trials, a cue (four dots surrounding a placeholder) was presented prior to the onset of the target. Like the target, the cue was defined by either onset or colour. Folk et al. (1992) found a typical cueing effect. That is, valid cues (which signalled the correct location of the to-be-presented target) resulted in faster responding to the target, while invalid cues (which signalled an incorrect location) resulted in slowed responding. However, cueing *only* occurred if the cue matched the properties of the target. Therefore, when the target was defined by onset, onset cues affected performance, and when the target was defined by colour, colour cues affected performance. This indicates that attentional orienting is driven, at least in part, by top-down attentional settings. Decades of subsequent research have supported this notion of contingent capture, such that salient cues only capture attention if they match a participant's top-down set (for a review, see Folk & Remington, 2010).

The vast majority of research on top-down attention has focussed on top-down sets for simple features (e.g., onsets and colours). However, there is some evidence that this contingent relationship between one's top-down set and the properties that capture attention extends to more complex stimuli, including faces (Barratt & Bundesen, 2012; Goodhew,

Kendall, Ferber, & Pratt, 2014; Wyble, Folk, & Potter, 2013). Barratt and Bundesen (2012) employed a flanker task in which participants responded to the expression of a central schematic face surrounded by distractor faces. Participants responded more slowly to happy face targets when they were flanked by negative faces compared with positive or neutral faces, indicating involuntary capture of attention by negative faces. By contrast, when the central target was a letter and not a face, participants' attention was not captured by negative face distractors in this way. This could reflect, in the latter case, that participants had a top-down set for a letter, and not the faces, thereby undermining the facial expression flanker's ability to capture their attention. This suggests that spatial attentional capture by faces is dependent on top-down attention. However, it is unclear what constituted participants' top-down set on trials when the target was a face. That is, participants could have had a top-down set for the generic category of "face" (or even a broader category, such as "complex object"), but equally it could have been "emotional face". There is no evidence that it was an emotion-specific top-down set (e.g., "happy face") as, if this set could be instantiated, we would expect that the negative face flankers would have been excluded from attentional capture. Here, therefore, we examined whether participants could adopt a top-down set for a specific emotional expression and, if so, whether this was driven by simple features or more holistic processing. Furthermore, we examined whether bottom-up processing of emotional expressions would interfere with the operation of any top-down set.

Selection History

In addition to top-down and bottom-up attention, a third mechanism of attention requires consideration: selection history. Selection history recognises that visual attentional capture can occur based on the individual's selection and reward history, which may differ from their current top-down goals. For example, Maljkovic and Nakayama (1994) found that when participants made speeded responses to a target defined by colour or spatial frequency

cues, RTs were faster when the same target repeated across two trials than when it switched. This difference occurred even when participants knew with 100% certainty the identity of the target on the upcoming trial, indicating that selection history can differ from one's current goal and guide attention. Some researchers have gone as far as to argue that participants cannot use top-down attention, and that all previous evidence for top-down effects can be explained purely by selection history (Theeuwes, 2013; Theeuwes, Reimann, & Mortier, 2006). More recent evidence, though, has shown that this is not the case; for example, when more complex arrays are employed, attention is guided by the participant's top-down set (Belopolsky & Awh, 2016).

Experimentally, there are several ways to distinguish between guidance by top-down attention and selection history. Firstly, throughout a block of trials, the target can be varied on a trial-by-trial basis (e.g., 200 trials in which the target is either red or green and is signalled at the start of each trial), and performance on trials for which the target type repeats across two trials versus when the target switches can be compared. Enhanced performance to the target on repeat trials, compared with switch trials, reveals an effect of selection history (also known an intertrial priming). A second method to measure selection history is to compare blocked versus mixed manipulations of the top-down set. That is, when the property that defines a target is fixed across a series of trials constituting an experimental block (e.g., search for a red target for 100 trials, and then a green target for 100 trials), then both top-down and selection-history processes are thought to contribute to attentional orienting. In contrast, in an intermixed block, the contribution of selection history is substantially diminished, providing a measure of top-down contributions (though note that, selection history effects can only be eliminated if the target never repeats across trials). Since top-down contributions can be considered as constant across the blocked versus mixed conditions, the impact of selection history can be gauged by comparing the blocked (selection history strong)

and mixed (selection history weak) designs (Awh et al., 2012; Büsel, Voracek, & Ansorge, 2018). Importantly, top-down attentional orienting can operate in mixed designs (Belopsolsky & Awh, 2016, Schoeberl, Goller, & Ansorge, 2019; Theeuwes, 2018). Since selection history has been found to play a critical role in attentional orienting to basic stimuli, it may also be central to understanding attentional orienting to faces. To our knowledge, no previous research has tested the relative contributions of top-down attention, bottom-up attention, and selection history to orienting to emotional faces.

Summary and Present Study

In sum, the extant empirical literature identifies three main categories of attentional guidance: top-down, bottom-up, and selection history. Previous studies on attentional orienting to facial expressions have purported that bottom-up mechanisms are responsible for attentional orienting. More recently, the importance of top-down attention toward faces has begun to be recognised but has not been systematically tested. The purpose of this study, therefore, was to test whether humans can instantiate a top-down set for a facial expression, and test whether this can guide attentional orienting in a way that is not accounted for by selection history and that overrides any bottom-up capture that may be present. These questions were assessed across four experiments.

Experiment 1

Experiment 1 used a modified dot-probe design with angry, happy, and neutral photographic face stimuli. In this task, participants were given a top-down cue (the words "Happy" or "Angry") to indicate the emotional expression they should orient towards. These cues could vary from trial-to-trial. Participants viewed two faces, one to the left and one to the right of fixation, and a probe letter replaced one of the faces. Valid trials were those in which the probe subsequently appeared in the same location as the cued face, whereas invalid trials were those in which the probe appeared in the opposite location to the cued face. The

majority of trials were valid (75%). Faster RTs to the probe in the location of the cued face (valid) compared with the opposite location (invalid) indicated that participants could rapidly orient to that face. This design was employed because it allowed measurement of attentional orienting towards faces via a widely-used and established measure of spatial attention (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; MacLeod, Mathews, & Tata, 1986). Moreover, since the same stimuli were presented on valid and invalid trials, cueing effects were not confounded by differences in responding when emotional stimuli were present compared with when not present (e.g., slowed responding or "freezing" can occur in the presence of threat but not in the presence of neutral-neutral trials).

It was predicted that, if participants instantiate a top-down set for happy and angry facial expressions, then they should show a cueing effect. By contrast, if attention toward either happy or angry facial expressions is driven by bottom-up attention, faster RTs will be found on trials in which the probe is presented in the locus of the emotional face, regardless of the cue. Finally, if there is an effect of selection history, a greater cueing effect will be found for trials in which the to-be-searched for facial expression repeats compared with when it switches.

Method

Participants. Forty-six participants were recruited via the Australian National University online sign-up portal, online Australian National University advertising portals, and flyers on campus. Ethical approval was provided by the Australian National University Human Research Ethics Committee (protocol number: 2014/534). Participants provided written, informed consent. Using G-Power[2], it was calculated that, to detect a small-medium effect size ($f = .15$) with 90% power, assuming a correlation of 0.6 between conditions, a minimum sample size of 42 was required (see Results section: after exclusions, 43 participants were included in the analyses). Restrictions were that participants were

Caucasian, aged 18-30 years, with normal or corrected-to-normal vision. Caucasian participants (self-reported) and Caucasian images were employed to control for the cross-race effect (for a review, see Young, Hugenberg, Bernstein, & Sacco, 2012), in which faces of other-race individuals can be processed differently to those of one's own race. Participants' ages ranged from 18 to 30 years ($M = 21.59$, $SD = 3.44$); 44 participants were right-handed, 2 were left-handed, 32 were female and 14 were male. The experiment took participants approximately 45 minutes to complete and each was offered either course credit or $10 cash.

Experimental stimuli and apparatus. This experiment was conducted in a dimly lit room. Stimuli were presented on a liquid crystal display (LCD) monitor running at a 60Hz refresh rate. Viewing distance was set with a chinrest at 44cm. Stimuli were programmed in Matlab using the Psychophysics Toolbox (Brainard, 1997). The background was set to black. Images of faces were taken from the NimStim database (Tottenham et al., 2009). The Nimstim database was selected because it has faces with closed-mouths. That is, closed-mouth expressions were chosen to eliminate attentional capture by teeth (Horstmann, Lipp, & Becker, 2012). In addition, the Nimstim database has been validated; recognition of facial expressions was 84% proportion correct for the closed-mouth angry faces, 92% for closed-mouth happy faces, and 91% for closed-mouth neutral faces (Tottenham et al., 2009).

The selected images consisted of Caucasian faces posing with closed-mouth neutral, happy, and angry expressions (Tottenham et al., 2009). Further exclusions were made due to the incorrect size of one of the images (model 1), confounding facial hair (model 31), and not having a closed-happy face image (model 27). Therefore 22 models (7 females and 15 males) were included, each providing happy, neutral, and angry expressions. During the experiment, each image subtended 9.4° x 12.1° of visual angle, with a gap of 9.4° of visual angle between the two images. On each trial, the word cue ("Happy" or "Angry") was presented in white,

size 18, Helvetica font. In addition, the probe letter ("E" or "F") was presented in white, size 18, Helvetica font.

Procedure. The conditions included happy-neutral, happy-angry, and angry-neutral face pairings. It was always either the happy or angry face that was cued, either validly (the target was in the cued location), or invalidly (the target was in the non-cued location). In this design, the main analysis was a 2 (target face: angry or happy) x 2 (distractor face: emotional or neutral) x 2 (cue: valid or invalid) factorial repeated-measures ANOVA.

An additional "incorrect cue" condition was included. Here, pairs of faces were happy-neutral or angry-neutral. The cue did not match the faces presented, and the probe appeared in the locus of the emotional expression. This condition was included as an additional test to ensure that participants' attention would not be captured by an emotional face that did not match their top-down set. An example of this type of trial would be when participants received a cue of "Angry" but were then presented with a happy-neutral face pair. The incorrect cue condition was not part of the main factorial analysis, since it did not have a corresponding valid condition, and was instead analysed separately.

Participants completed a demographic questionnaire and then the experimental computer task. The demographic questions assessed age, gender, handedness, and cultural/racial identity. The computer task included 50 practice trials and then 360 experimental trials with 8 rest breaks. Trials were randomised throughout the experiment with the constraint that the cued face was predictive of the probe location on 75% of trials. A break-down of trials was as follows: 135 trials were either valid happy-neutral (cued happy) or angry-neutral (cued angry) face pairs; 135 trials were either valid happy-angry (either validly cued happy or angry) face pairs; 30 trials were invalid happy-neutral or angry-neutral face pairs; 30 trials were invalid happy-angry face pairs; and 30 trials were incorrectly cued happy-neutral or angry-neutral face pairs. The cue (happy or angry) within each condition

was randomised, such that an approximately equal number of trials were included for each cue. On each trial, the two images presented were of the same model so that they were matched for facial properties. The identity of the face on each trial was randomly selected with the constraint that the same proportion of male and female images were included in each condition.

On each trial, participants were presented with a blank black screen for 500ms with a centrally presented fixation cross (see Figure 1). The cue word ("Angry" or "Happy"), written in the centre of the screen in white, was then presented for 1000ms and was followed by a blank screen with a fixation cross for 1000ms. Participants were instructed to orient their attention to the cued facial expression when it appeared. The faces were presented for 200ms, one to the left and one to the right of fixation. This timing was selected with the intention of providing sufficient time for participants to shift their attention toward one of the face locations without being able to make multiple saccades. After face offset, a probe (the letter "E" or "F") was presented for 300ms in the locus of one of the faces. Participants were asked to indicate the identity of the probe (an E or F), with a keyboard press, as accurately and quickly as possible. Once participants made a keyboard press, the next trial began.

It is important to note that this was a speeded task, and participants were *not* instructed to withhold their response during the probe. Instead, the probe duration was sufficiently short that meaningful responses would unlikely be made during this time (indeed 200-300ms is a commonly-employed criterion for excluding artifactual RTs from identification responses). The advantage of having a time-limited probe duration (rather than shown until response) is that it more strongly compels attention to be in the probe location, and a greater cost is incurred if it is not (since it subsequently disappears). Thus, the design is most sensitive to measuring the spatial allocation of attention. Converging evidence for this is

provided by previous research that has used such a design and found robust attentional biases

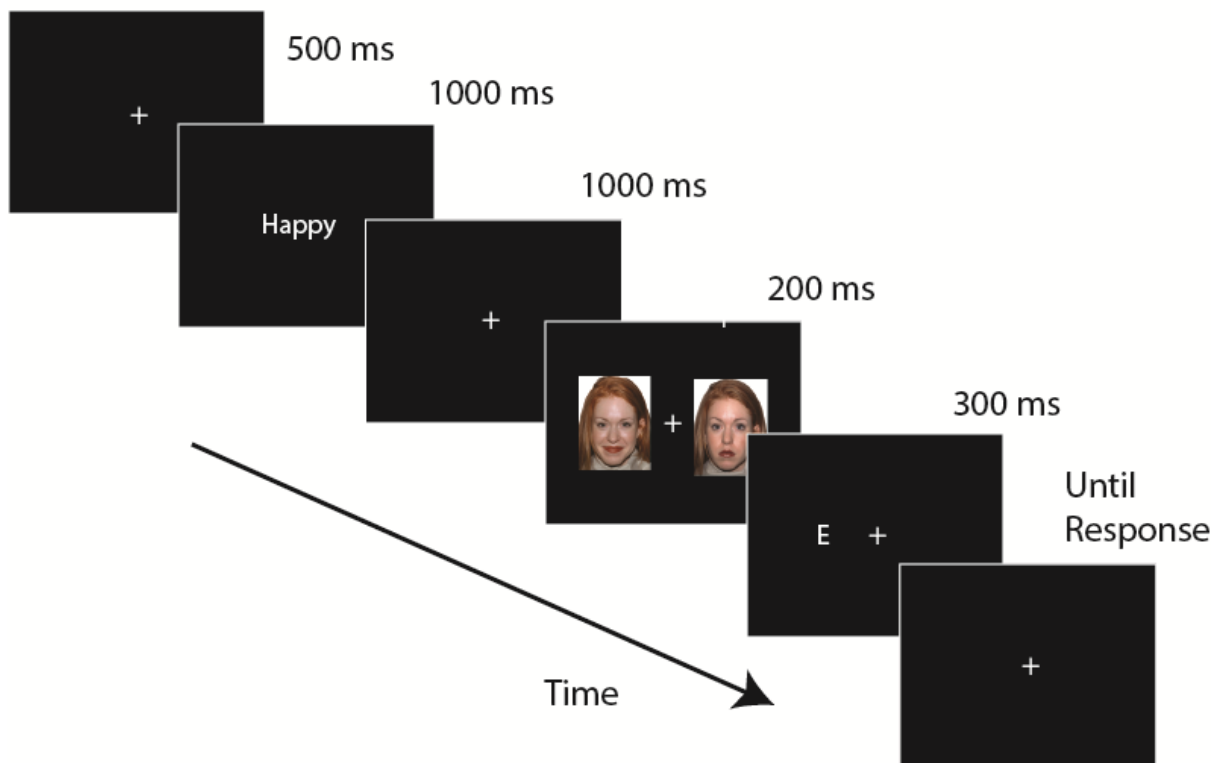(see e.g., Cox, Christensen, & Goodhew, 2017).

Figure 1. Schematic of a trial. On this trial, the cue is "Happy", which is followed by a

happy-neutral face pair, and the probe is validly presented in the locus of the happy face.

Note that participants were instructed to respond as quickly and accurately as possible as

soon as they could identify the probe.

Results and Discussion

  Two participants' data were excluded due to chance-level accuracy (49.9% and

49.7%) on the probe detection task and one participant's data were excluded due to unusually

slow RT (z-score > 3.29). Therefore, data from 43 participants were included in the final

analyses. Trials in which participants made an invalid key press (i.e., hit a key that was not

"E" or "F") or responded quicker than 100ms or slower than 2.5 standard deviations from

their average RT were also excluded. This accounted for 2.3% of participants' data. Accuracy

and RTs for the correct trials were then calculated. Mean accuracy on the probe task was 91% (SD = 6.6%) and mean RT was 674.9ms (SD = 103.7ms). Full raw data for all experiments are available here: [https://osf.io/8zt2a/])[3].

Cueing effect. Although the RT data are of primary interest to the analyses, accuracy was first assessed to test for any potential speed-accuracy trade-off. To assess accuracy, a 2 (target face: happy or angry) x 2 (distractor face: neutral or emotional) x 2 (cue: valid or invalid) repeated-measures ANOVA was conducted. No significant effects were found (ps ≥ .135).

To test whether participants demonstrated a cueing effect, and if the cueing effect differed as a function of the target and distractor facial expressions, a 2 (target face: happy or angry) x 2 (distractor face: neutral or emotional) x 2 (cue: valid or invalid) repeated-measures ANOVA on RT was conducted. A significant effect was found for type of distractor face, $F(1, 42) = 6.04$, $p = .018$, $\eta^2 = .126$, such that participants were faster at responding to the probe when the distractor was neutral (i.e., happy-neutral and angry-neutral face trials) ($M = 670.9$ms), compared with when the distractor was emotional (i.e., happy-angry face trials) ($M = 679.4$ms). This result was qualified by an interaction between target face and distractor face, which was trending toward significance, $F(1, 42) = 3.84$, $p = .057$, $\eta^2 = .084$. Further analyses revealed that participants were significantly faster at responding to the probe when a happy-target face was paired with a neutral face ($M = 664.9$ms) compared to when paired with an angry face ($M = 679.6$ms), $F(1, 42) = 9.48$, $p = .004$, $\eta^2 = .184$. By comparison, RTs to the probe did not significantly differ when the target angry face was paired with a neutral face compared with when paired with a happy face, $F(1, 42) = .27$, $p = .607$, $\eta^2 = .006$.

However, the interactions between distractor face and cue validity ($p = .945$) and among target face, distractor face, and cue validity ($p = .777$) were not significant, indicating that the cueing effects were not altered by the type of distractor face or target face. This

suggests that, regardless of whether participants were responding to the validly or invalidly

presented probes, participants had slower RTs in the presence of an angry face (either when

the angry face was the target or the distractor face). This may reflect a freezing effect in the

presence of threat, which leads to slowed responses when angry faces were present (Roelofs,

Hagenaars, & Stins, 2010).

Most importantly, a significant cueing effect was found, $F(1, 42) = 7.36$, $p = .010$, $\eta^2 = .149$, as participants were faster on valid trials ($M = 666.4$ms) compared with invalid trials ($M = 683.9$ms) (see Figure 2). No other significant effects were found ($ps \geq .107$). In summary, these results indicate that participants could reliably shift their attention toward both the happy and angry face in response to top-down cues[4].
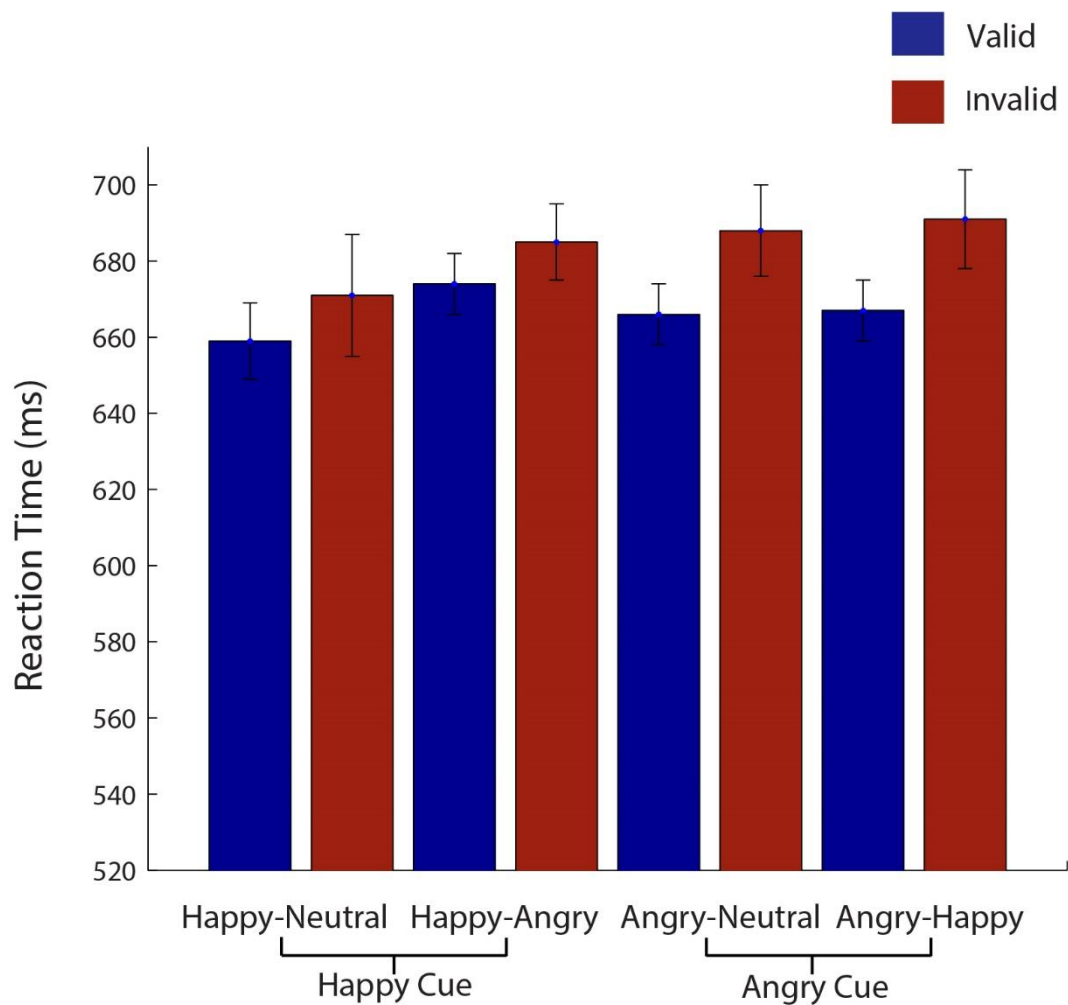
Figure 2. Reaction data for Experiment 1. Within-subjects 95% confidence intervals were calculated using Morey's (2008) correction for within-subject designs.

Incorrect-cue trials. For incorrect-cue trials, the cue could either be "Happy" in the context of an angry-neutral face pair or "Angry" in the context of a happy-neutral face pair. This condition assessed whether participants would only orient to an emotional face when it matched their top-down set or whether bottom-up emotional capture of attention would occur toward the un-cued emotional face. Specifically, quick responding to the probe replacing these un-cued emotional faces would indicate bottom-up attentional capture. Conversely, if responding was slower than the valid conditions, it would indicate that participants were able to inhibit attentional capture by these emotional faces. To test this, the valid happy-neutral and angry-neutral data, which were previously used in the main analysis, were also analysed here.

A 2 (target face: happy or angry) x 2 (cue: valid or incorrect-cue) ANOVA was conducted. A significant cueing effect was found, $F(1, 42) = 5.03$, $p = .030$, $\eta^2 = .107$, with participants responding faster on valid trials ($M = 662.1$ms) compared with invalid incorrect-cue trials ($M = 673.7$ms). No other significant effects were found ($ps \geq .294$). This suggests that the correct cue provided an advantage in responding to the probes replacing emotional faces over and above responses following incorrect cues. In other words, top-down attention provided benefit over and above the possible effects of bottom-up attentional orienting to emotional faces.

Selection history effect. To measure the effect of selection history, cueing effects were compared between repeat and switch trials. Switch trials were those in which participants were asked to attend towards an expression that differed from the one attended to on the previous trial whereas, on repeat trials, participants were asked to attend to the same expression as on the previous trial. Trials were only included in these analyses if the previous

trial was valid as, for valid trials, shifting toward the cued face was "rewarded" as the probe

was presented in the locus of the cued face.

To test the effect of selection history, a 2 (target face: happy or angry) x 2 (distractor

face: neutral or emotional) x 2 (cue: valid or invalid) x 2 (repetition: repeat or switch)

ANOVA was conducted. The interaction between cue and repetition was not significant ($p =$

.737), as the cueing effect for repeat trials ($M_{valid} = 665.6$ms; $M_{invalid} = 682.2$ms) was not

significantly larger than the cueing effect for switch trials ($M_{valid} = 667.8$ms; $M_{invalid} =$

681.9ms). In other words, there was no evidence for selection history effects in the present

study[5]. Since the results of the selection history analysis failed to reject the null hypothesis,

the null hypothesis statistical testing was supplemented with Bayesian analysis performed in

JASP with the default priors. This revealed that the BF_{10} factor $= 0.12$ for the interaction

between cue validity and interaction (compared against a model that contains all effects

except this interaction). This quantifies the ratio of evidence in favour of the alternative

versus the null hypothesis for the interaction (Jarosz & Wiley, 2014), and therefore given that

it was <1, there was greater evidence *in favour* of the null hypothesis in this experiment. This

converges with the results of the hypothesis testing.

To summarise, Experiment 1 results suggest that participants could employ top-down

attention to orient toward both angry and happy facial expressions, as evidenced by

significant cueing effects in response to top-down cues. In addition, cueing effects did not

significantly differ between trials when an emotional face (happy or angry) was paired with a

neutral face or when happy and angry faces were paired together. Furthermore, one of the

conditions in this study (incorrect-cue trials) cued an expression that was not subsequently

presented. Responses in this condition were slower than responses following correct cues.

This indicates that participants used the top-down cue to reliably orient attention to the cued

face. Finally, no evidence for selection history effects was found. These results provide

support for the notion that top-down control can be used to guide attention to facial

expressions. While a contribution of bottom-up attentional capture cannot be ruled out, no

evidence for bottom-up attentional capture was found (this issue was addressed explicitly in

Experiment 3, where it was confirmed that bottom-up attentional capture did not play a role).

This, in conjunction with the evidence that valid cues were providing orienting advantage

over incorrect cues on incorrectly cued trials supports the conclusion that the present results

do indeed reflect top-down attentional orienting.

Experiment 2

Although Experiment 1 found that participants could shift their attention toward the

cued happy and angry faces, the specific properties of the faces used to do so remain an open

question. That is, do participants rely on specific facial features (e.g., the mouth or eyes) to

shift attention to cued happy or angry faces or do they form a holistic top-down set for those

facial expressions? Research has found that inverting a face disrupts the ability to process the

face holistically and induces a more piecemeal processing of separate face parts (Mondloch

& Maurer, 2008; Searcy & Bartlett, 1996; Yin, 1969). Experiment 2, therefore, tested

participants' ability to shift attention to cued happy and angry faces on the dot-probe task

with *inverted* (i.e., upside-down) faces. If participants are able to shift attention to the cued

expressions, as in Experiment 1, it would indicate that top-down attentional sets for facial

expressions can be instantiated by specific facial features (e.g., upturned lips could drive

attention to happy faces). If, however, participants cannot shift attention to the cued

expression, it would indicate that top-down orienting to facial expressions requires holistic

processing.

Participants. Forty-seven participants were recruited via the Australian National

University Online sign-up portal. Participants' ages ranged from 18 to 29 years ($M = 20.21$,

$SD = 2.35$), 46 participants were right-handed and 1 was left-handed, 19 participants were

male, 27 were females, and 1 participant did not disclose their gender. Participation took approximately 10 minutes. Participants completed Experiment 3 and then Experiment 2 and were offered 45 minutes of course credit.

Experimental stimuli, apparatus, and procedure. Experiment 2 replicated Experiment 1, with two exceptions: (1) the faces were inverted, and (2) the emotion-neutral (incorrect cue) condition was excluded since it was not essential to test the hypothesis. Furthermore, since participants were also required to complete Experiment 3 with Experiment 2, it was helpful to reduce the trial numbers so that participation time could be kept to 45 minutes to reduce fatigue effects. The following face pairs were used: happy-neutral, angry-neutral, and happy-angry faces. The probe could either be presented in the cued or un-cued location. This experiment consisted of 12 practice trials and 320 experimental trials with 6 rest breaks.

Results and Discussion

One participant's data were excluded due to pressing the wrong keys (100% invalid key press). A further two participants' data were excluded as they did not meet criteria for being Caucasian and one additional participant's data were excluded as the RT constituted an outlier (z-score > 3.29). Therefore, 43 participants' data were included in further analyses. Trials in which participants made an invalid key press (i.e., hit a key that was not "E" or "F") or responded quicker than 100ms or slower than 2.5 standard deviations from their average RT were also excluded. This accounted for an average of 2.2% of participants' data. Accuracy and RTs for the correct trials were then calculated. Mean accuracy on the probe task was 94% ($SD = 4.7\%$) and mean RT was 643.2 ($SD = 93.6ms$).

Cueing effect. Accuracy was analysed with a 2 (target face: happy or angry) x 2 (distractor face: neutral or emotional) x 2 (cue: valid or invalid) repeated-measures ANOVA. No significant effects were found ($ps \geq .148$). To test whether participants could shift

attention to the cued expressions, RT was analysed with a 2 (target face: happy or angry) x 2 (distractor face: neutral or emotional) x 2 (cue: valid or invalid) repeated-measures ANOVA. No significant effects were found ($ps \geq .085$), indicating that participants could not reliably shift attention to the cued expression and that the effects did not significantly differ depending on target facial expression and type of paired expression. Specifically, the cueing effect was clearly non-significant, $F(1, 42) = 1.78$, $p = .190$, $\eta^2 = .041$), with similar RTs on valid trials ($M = 641.1$ms) compared with invalid trials ($M = 645.4$ms) (see Figure 3). Since the results of this experiment failed to reject the null hypothesis, the null hypothesis statistical testing was supplemented with a Bayesian analysis. This revealed that the BF_{10} factor for the effect of cue validity was .34. This means that there was greater evidence *in favour* of the null hypothesis in this experiment, thereby converging with the results of the hypothesis testing.
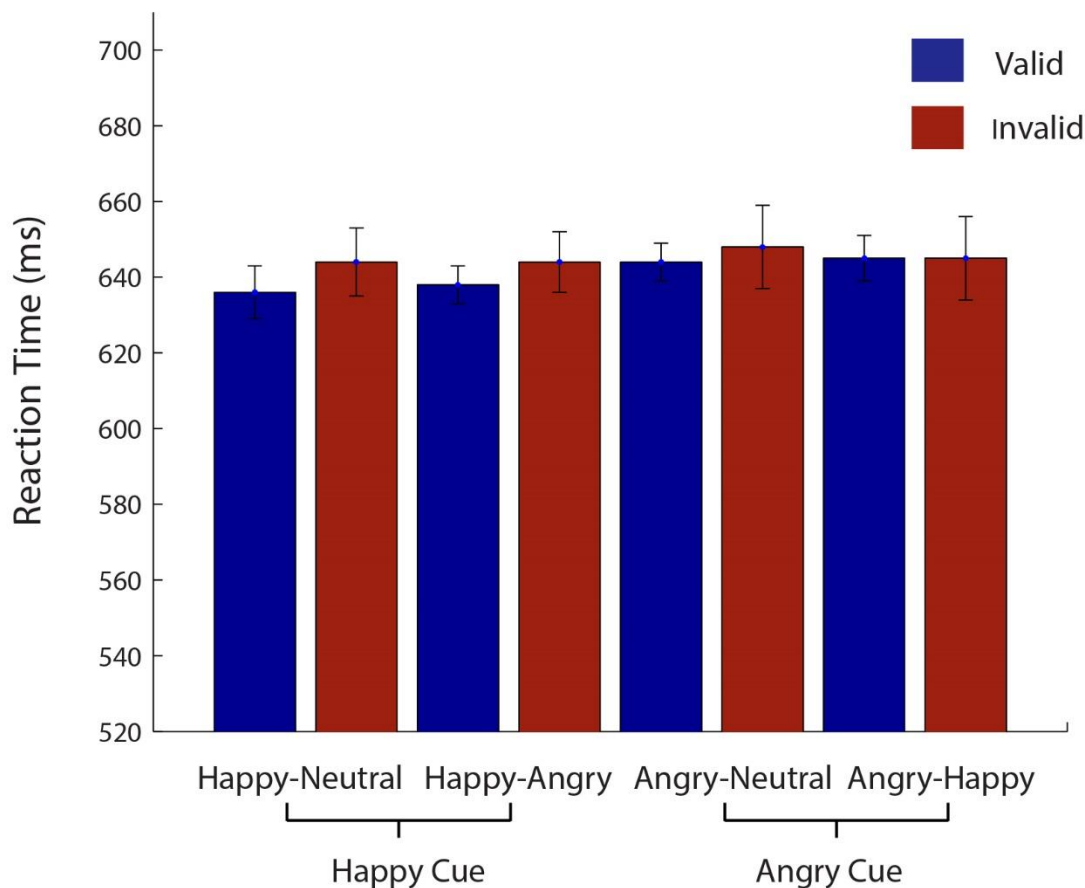
Figure 3. Reaction data for Experiment 2. Within-subjects 95% confidence intervals were calculated using Morey's (2008) correction for within-subject designs.

To summarise, Experiment 2 failed to reveal any evidence that participants were able to use top-down attention to orient to inverted emotional faces. This was evidenced by a non-significant cueing effect. In conjunction with Experiment 1, where there was reliable orienting toward facial expressions, this indicates that top-down attention toward facial expressions is based on perceiving the face holistically rather than searching for individual face parts. Interestingly, the overall RTs from Experiment 2 were faster compared with the RTs in Experiment 1. One possibility is that, because participants could not reliably orient to the inverted face, they tended to disregard the faces and instead relied on bottom-up capture of attention by the onset of the probe to perform the task. This could have led to slightly faster probe detection than in Experiment 2. Critically, however, faster RTs were not found

on valid, relative to invalid, trials indicating that participants could not reliably orient to the cued face when the faces were inverted.

Experiment 3

The results of Experiment 1 demonstrated that participants could shift attention to both happy and angry faces, and that the cueing effects did not significantly differ between the two expressions. However, this study did not test participants' baseline bottom-up capture of attention by happy and angry faces. This is an important consideration because participants may have bottom-up biases to happy and/or angry faces. In determining these bottom-up biases, we can then understand how top-down attention may override or compete with bottom-up attention.

The incorrect-cue condition partly measured bottom-up capture of attention, as it tested participants' ability to ignore irrelevant emotional stimuli. However, in this condition, participants were still given a top-down set ("Happy" or "Angry"). To provide a purer measure of bottom-up attention, no top-down cues should be provided. To this end, Experiment 3 utilised a modified dot-probe task of happy-neutral, angry-neutral, and happy-angry paired trials and participants did not receive any instructions of which faces to attend. In addition, the probe was equally likely to follow angry, happy, or neutral faces and so there were no top-down effects regarding expectations of where the probe was most likely to appear. The following experiment, therefore, tested whether participants have an attentional bias toward happy or angry facial expressions.

Method

Participants. The sample from Experiment 2 also completed Experiment 3.

Experimental stimuli, apparatus, and procedure. Each trial consisted of 500ms of blank screen with a fixation cross, then the paired faces were presented for 200ms, and was followed by 300ms of probe presentation. Participants were then presented with a blank

screen which remained visible until they made a response. Crucially, however, they were

instructed to identify the probe as quickly and accurately as possible as soon as they saw it.

As with Experiment 1, paired faces could be happy-neutral, angry-neutral, or happy-angry.

The six trial types were: (1) happy-neutral, probe follows happy, (2) happy-neutral, probe

follows neutral, (3) angry-neutral, probe follows angry, (4) angry-neutral, probe follows

neutral, (5) happy-angry, probe follows happy, and (6) happy-angry, probe follows angry.

This task consisted of 12 practice trials and then 120 experimental trials, with two rest breaks.

The presentation of the six trial types was randomised throughout the experiment with the

restriction that there were 20 trials for each condition. All other stimuli and apparatus details

were identical to Experiment 1.

Results and Discussion

As with Experiment 2, one participant's data were excluded due to 100% invalid key

presses and a further two participant's data were excluded as they did not meet criteria for

being Caucasian. However, the participant's data which constituted an outlier in Experiment

2 was not an outlier in Experiment 3. Therefore, data from 44 participants were included in

the analyses of Experiment 3. Trials in which participants made an invalid key press (i.e., hit

a key that was not "E" or "F") or responded quicker than 100ms or slower than 2.5 standard

deviations from their average RT were also excluded. This accounted for 2% of participants'

data. Accuracy and RTs for the correct trials were then calculated. Mean accuracy on the

probe task was 92.9% ($SD = 5.1\%$) and mean RT was 576.2ms ($SD = 59.7$ms).

Accuracy was first assessed across the six trial types. Using a 6 (trial type) repeated-

measures ANOVA, no significant differences in accuracy were found, $F(5, 215) = .39$, $p =$

.855, $\eta^2 = .009$. To test whether participants' RT differed across the six types of trials

depending on paired facial expression and probe location, a 6 (trial type) repeated-measures

ANOVA was conducted. No significant differences were found between the trial types, $F(5,$

215) = 1.07, p = .377, η^2 = .024 (see Figure 4). This indicates that, with an unselected sample, there are no preferences to attending to happy, angry or neutral facial expressions.
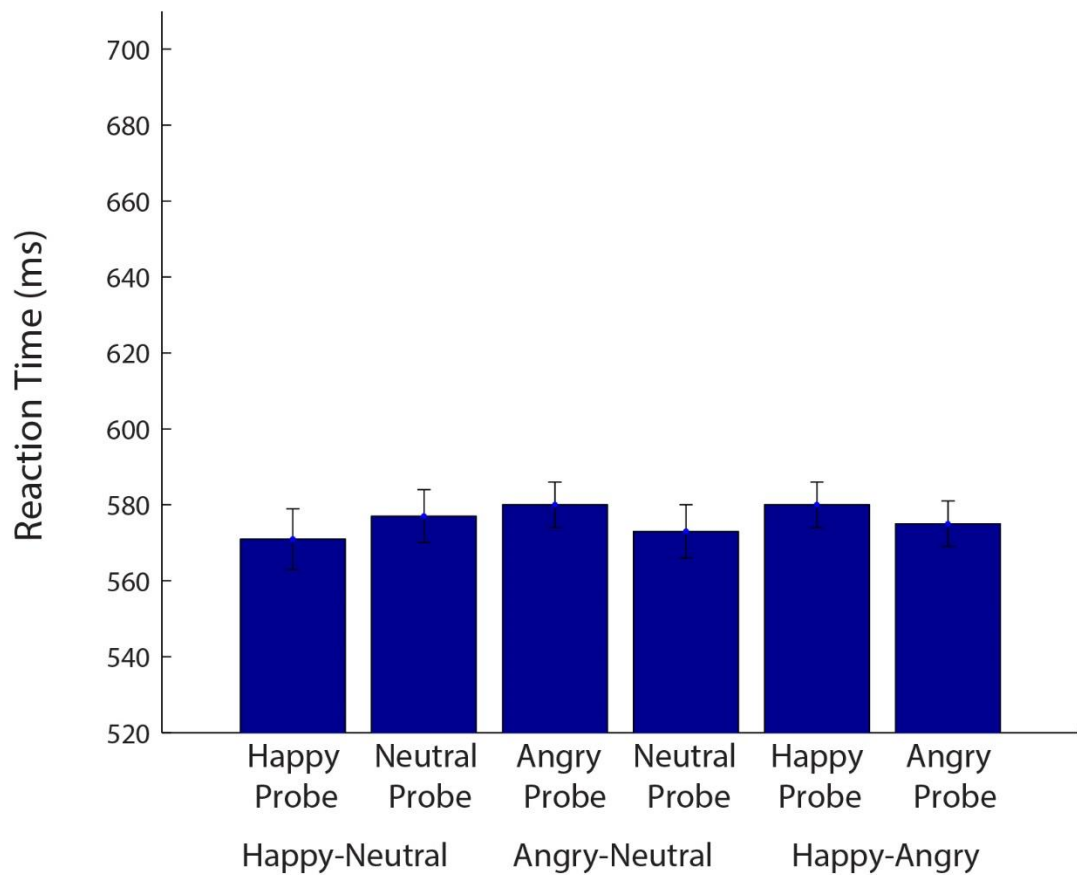

Figure 4. Reaction data for Experiment 3. Within-subjects 95% confidence intervals were calculated using Morey's (2008) correction for within-subject designs.

Once again, since the results of this experiment failed to reject the null hypothesis, the null hypothesis statistical testing was supplemented with Bayesian analysis. This revealed that the BF_{10} for the effect of trial type was .046. This means that there was greater evidence *in favour* of the null hypothesis for the effect of trial type in this experiment. Therefore, this converges with the results of the hypothesis testing.

In summary, Experiment 3 found no evidence for a bias toward happy, angry, or neutral facial expressions. Interestingly, the overall mean RTs were shorter in Experiment 3, compared with the RTs from Experiment 1 and Experiment 2. This indicates that not having a

top-down cue for an emotional expression and, instead, relying on bottom-up attentional capture by the probe, resulted in faster RTs to the probe. Crucially, however, the results of Experiment 3 reveal that, in the absence of a top-down set, participants did not have a bottom-up bias to emotional faces.

<div align="center">**Experiment 4**</div>

Experiment 1 did not support selection history as a mechanism of attention to specific facial expressions. Experiment 4 was conducted to examine the effect of selection history more fully. Specifically, it is possible that selection history effects require more than one repeated trial to result in increased cueing effects. For this reason, Experiment 4 compared a mixed condition, in which the cued face (happy or angry) changed randomly throughout the trial block, with a blocked condition in which participants attended to the same facial expression throughout a block of trials. If cueing effects are larger in the blocked condition compared with the mixed condition, it would indicate that selection history increases participants' ability to rapidly shift attention to emotional faces. However, if the cueing effect does not differ between the blocked and mixed conditions, this would lend further support to the conclusion that selection history is not a major factor in guiding attention to emotional faces.

Method

Participants. Forty-six participants were recruited via the Australian National University online sign-up portal in exchange for course credit. Participants' ages ranged from 18 to 24 years ($M = 19.76$, $SD = 1.59$), 39 participants were right-handed, 6 were left-handed, and 1 identified as ambidextrous, 26 participants were females, 19 were male, and 1 did not disclose their gender. Participation took approximately 30 minutes.

Experimental stimuli, apparatus, and procedure. The stimulus presentation parameters were identical to Experiment 1. However, this experiment consisted of three tasks, which were counterbalanced across participants:

1. Happy block: Participants were instructed to shift their attention toward the happy facial expression. Trials consisted of happy-neutral and happy-angry face pairs. This block consisted of 12 practice trials, and then 160 experimental trials with 7 rest breaks.

2. Angry block: Participants were instructed to shift their attention toward the angry facial expression. Trials consisted of angry-neutral and angry-happy face pairs. This block consisted of 12 practice trials, and then 160 experimental trials with 7 rest breaks.

3. Mixed block: Happy and angry cue trials could randomly change each trial. Face pairs could be happy-neutral, angry-neutral, and happy-angry. This block consisted of 12 practice trials and then 320 experimental trials with 7 rest breaks.

Trials types were randomised throughout the experiment with the constraint that the cued face was predictive of the probe location on 75% of trials. An equal number of cued happy and angry trials were included and an equal number of angry-neutral, happy-neutral, and angry-happy face pair trials were included.
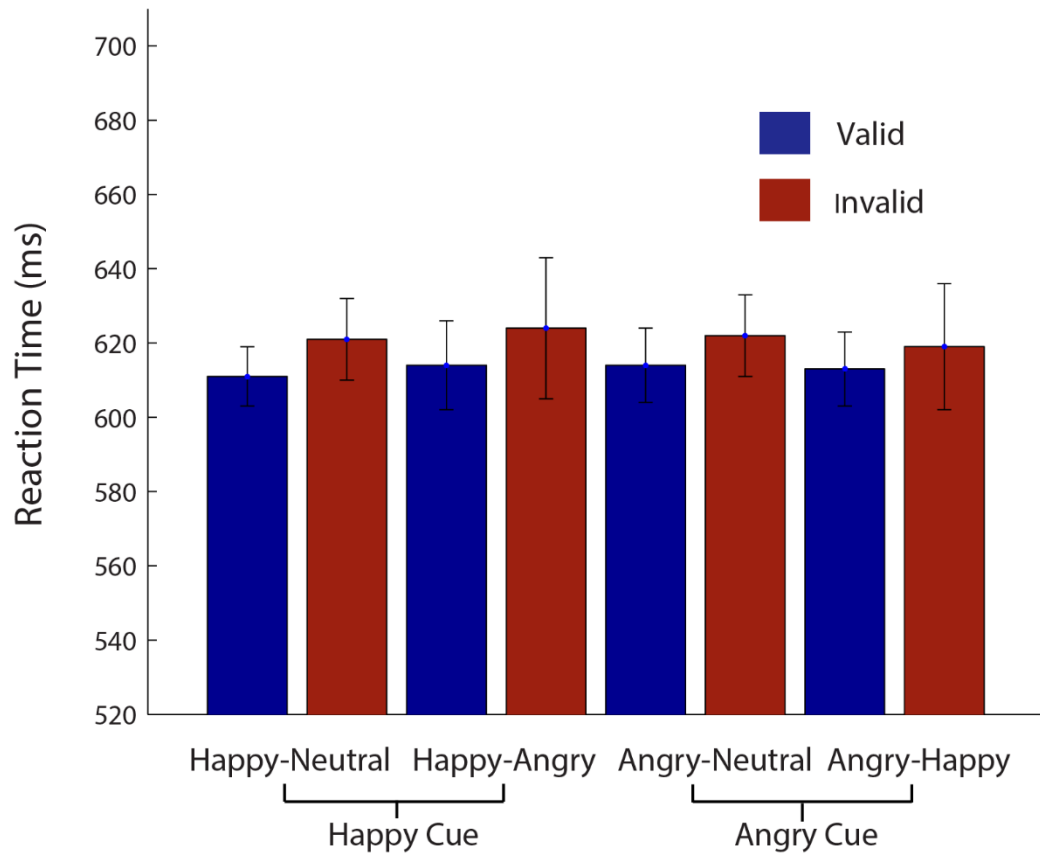
Results and Discussion

Data exclusions were as follows: three participants' data as they did not meet criteria for being Caucasian, two participants' data for chance accuracy, one participant's data for not having normal or corrected-to-normal vision, and four participants' data for outlier RTs (z score > 3.29). Therefore, 36 participants' data were included in the study. Next, 2.2% of data were excluded for invalid key presses and RTs that were quicker than 100ms or slower than 2.5 *SD* from each participant's mean RT. Accuracy and RTs for the correct trials were then

calculated. Mean accuracy on the probe task was 91.5% (SD = 5.1%) and mean RT was 600.4 ms (SD = 57.2 ms).

Cueing effect. Accuracy was analysed with a 2 (condition: mixed or blocked) x 2 (target face: happy or angry) x 2 (distractor face: neutral or emotional) x 2 (cue: valid or invalid) repeated-measures ANOVA. A significant effect of condition was found, $F(1, 35)$ = 23.10, p < .001, η^2 = .398, with higher accuracy in the mixed condition (M = 93.1%) compared with the blocked condition $(M$ = 89.7%). No other significant effects were found (ps ≥ .064).

To test whether blocking the top-down cue increased the cueing effect, a 2 (condition: mixed or blocked) x 2 (target face: happy or angry) x 2 (distractor face: neutral or emotional) x 2 (cue: valid or invalid) repeated-measures ANOVA on RT was conducted. A significant effect was found for cue validity, $F(1, 35)$ = 6.96, p = .012, η^2 = .166, with faster RTs on valid (M = 597.3ms) compared with invalid trials (M = 603.6ms). Condition was also significant, $F(1, 35)$ = 14.23, p = .001, η^2 = .289, with faster RTs in the blocked condition (M = 583.5ms) compared with the mixed condition (M = 617.4ms). However, although participants were generally faster in the blocked condition, the interaction between condition and cue validity was not significant, $F(1, 35)$ = .88, p = .354, η^2 = .025, indicating that the cueing effect did not significantly increase when the target face was blocked (see Figure 5). Therefore, no evidence for selection history was found. No other significant effects were found ($ps.$ ≥ .191)[6]. Convergent evidence arose from the Bayesian analysis, which indicated evidence in favour of the null hypothesis for the interaction between condition and cue validity (compared against a model with all effects except this interaction), (BF_{10} = .57).
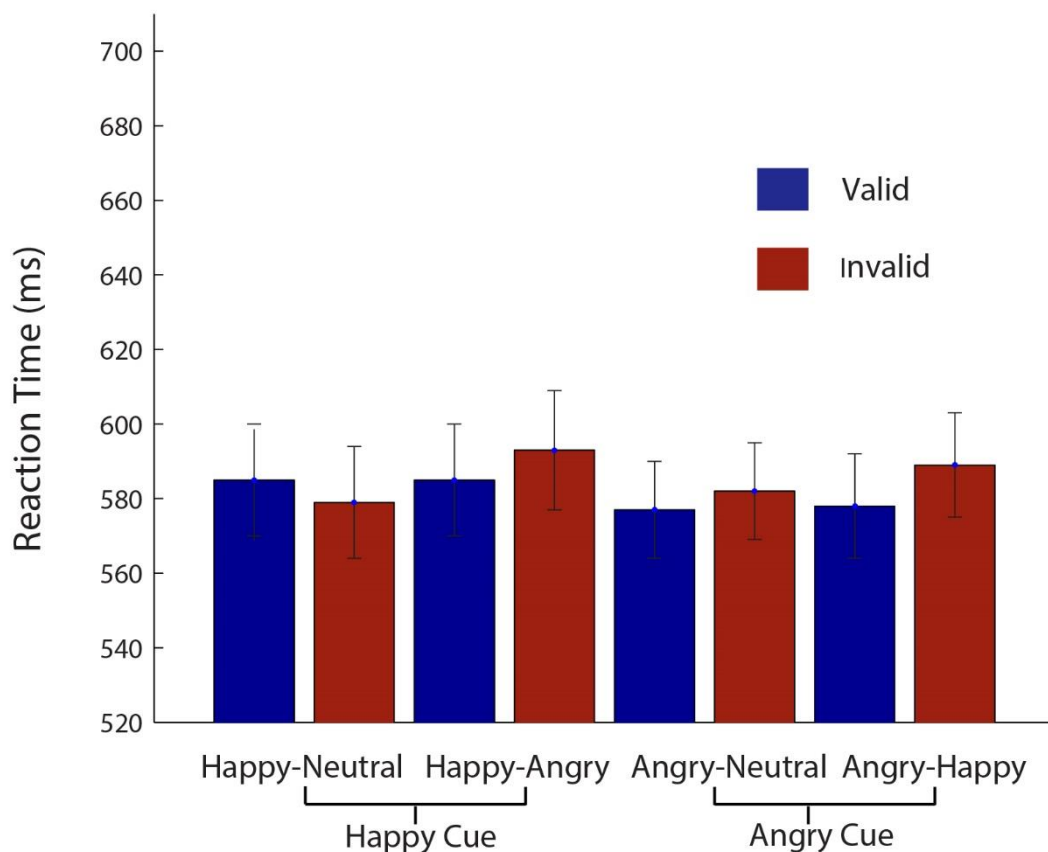
Figure 5. Reaction data for Experiment 4 for each cued face and face-pair separately.

Mixed condition is presented TOP and blocked condition is presented BOTTOM. Within-

subjects 95% confidence intervals were calculated using Morey's (2008) correction for

within-subject designs.

In summary, Experiment 4 compared conditions in which participants were instructed

to attend to one facial expression in a block of trials (blocked condition) compared with when

the cued face randomly changed across trials (mixed condition). An overall speed-accuracy

trade-off effect was revealed, with faster RTs but reduced accuracy for the blocked condition.

Critically, however, there was a reliable cueing effect, with faster RTs on valid versus invalid

trials, indicative of the fact that participants shifted their attention to the cued face. This

cueing effect was evident in both block conditions and hence uncontaminated by the speed-

accuracy trade-off. Critically, no effect of selection history was found as the cueing effect did

not significantly differ between the mixed and blocked conditions. That is, repeated practice

of attending to the same facial expression did not improve participants' ability to attend to the

cued face compared with the opposite location.

   Gender effect. Although not directly relevant to the hypotheses of the current study,

data from each of the four experiments were also analysed to test if the gender of the faces

employed and the gender of the participants affected results (e.g., Becker, Kenrick, Neuberg,

Blackwell, & Smith, 2007; Hess, Blairy, & Kleck, 1997; Wells, Gillespie, & Rotshtein,

2016). No consistent results were found across experiments. For example, Experiment 1

found slightly faster RTs to the probe on trials with male, happy target faces paired with a

neutral face compared with an angry-neutral face-pair ($p = .004$). In addition, on happy-

neutral and angry-neutral face-pair trials, female participants had a larger cue validity effect

compared with male participants ($p = .05$). However, there were only 14 male participants,

which is a relatively small sample size and, therefore, these results should be interpreted with

caution. No other significant main effects or interactions were found for face gender ($ps \geq$

.219) or participant gender ($ps \geq .124$). Next, Experiment 2 found that female participants,

but not male participants, displayed significantly ($p = .031$) faster RTs on happy-target trials

($M = 636.9$ms) compared with angry-target trials ($M = 647$ms). No other significant main or

interactive effects were found for face gender ($ps \geq .160$) or participant gender ($ps \geq .174$).

Experiment 3 found no significant main or interactive effects for face gender ($p \geq. 533$) or

participant gender ($p \geq. 574$). Similarly Experiment 4 found no significant main or interactive

effects of face gender ($p = .346$) or participant gender ($p = .278$). Altogether, the

inconsistency in the reliability of these results across the experiments suggest that the gender

of the faces or participants were not important factors in this study.

General Discussion

Across four experiments, the current study used a modified dot-probe design to test the relative contributions of top-down attention, bottom-up attention, and selection history mechanisms for shifting attention to emotional faces. The major finding of the current study was that participants could shift their attention in a top-down manner toward both angry and happy facial expressions, as evidenced by significant cueing effects in response to top-down cues. It was found that an unselected sample of participants did not have a bias toward either happy, angry, or neutral facial expressions (Experiment 3). Instead, when given a cue to attend to happy or angry faces, participants were able to rapidly deploy attention to those cued expressions using top-down attention (Experiment 1). Experiment 2 suggested that holistic rather than feature-based processing of the face was implicated in this top-down set, as the cueing effect was eliminated when the faces were inverted. Furthermore, the evidence from both Experiments 1 and 4 implies that this was driven by top-down, rather than selection history effects, as target repetition did not benefit attentional orienting.

These findings challenge the prevailing notion that attentional orienting to facial expressions in driven by bottom-up factors. Instead, here, top-down processes were clearly the dominant contributing factor. While we did not find any evidence for bottom-up processes, this does not preclude the possibility that they can, at times, contribute to attentional orienting to faces, in certain contexts or perhaps for certain individuals (such as those with higher levels of social anxiety). What it does mean, however, is that top-down attentional processes play a key role in how attention is applied to human facial expressions. This would seem to belie the notion of a "fear-relevant" module in the brain (e.g. Öhman & Mineka, 2001) that process threatening facial expressions in a way that is impervious to top-down attentional processes.

The importance of top-down attention in orienting toward facial expressions demonstrated in the present study is consistent with the conclusions of other recent research. Delchau, Christensen, O'Kearney, and Goodhew (in press) found that the attentional bias to threatening faces that is present in individuals with high levels of social anxiety was abolished under a working memory load. If the bias were truly automatic or bottom-up, it would be resistant to such contextual factors. It was not, and instead it was affected by performing a task that is thought to draw on shared resources with top-down attention, thereby implicating top-down processes in the bias itself. This is consistent with the conclusions of the present study.

An important methodological strength of the current experiments was the inclusion of the happy-angry face pairs. The fact that reliable cueing occurred on those trials rules out alternative, non-top-down explanations for the orienting, such as bottom-up capture by emotional faces. That is, in theory, bottom-up capture by emotional faces could have led to cueing effects for the angry-neutral and happy-neutral pairs. However, this possibility is refuted by the fact that participants were also able to orient their attention to either the happy or angry face when it was paired with an emotional face distractor. A bias toward emotional faces in general would not have produced reliable attentional orienting effects for such trials. In addition, if only happy-neutral and angry-neutral trials had been included in the study, participants could have developed a top-down set for emotional faces, or selection history could have emerged producing a bias toward any emotional face, since the probe was more likely to appear behind the emotional face compared with the neutral face. However, reliable cueing effects were found on happy-angry trials, irrespective of whether the happy or angry face was the target. This indicates that participants were flexibly able to switch, on a trial-by-trial basis, between orienting to happy and angry faces. Further evidence for this is the fact that participants were reliably faster to respond to a validly cued probe on happy-neutral and

angry-neutral trials compared with when the probe followed an incorrectly cued emotional face (e.g., a happy cue preceded an angry-neutral face pair). Thus, the content of the top-down set had to include at a *minimum* the valence of the particular expression (e.g. positive / negative), if not the specific expression itself (happy / angry). Future research could include additional expressions (e.g. sadness) to distinguish between valence and specific-expression-based content of the top-down set.

Previous research exploring top-down attentional orienting has largely employed basic visual stimuli, such as colour, shape, and motion cues, (for a review, see Folk & Remington, 2010). However, Wyble et al. (2013) demonstrated contingent capture effects for superordinate-level visual stimuli, such as sports equipment, food, and animals. The results of the current study indicate that, in addition to these types of stimuli, individuals can orient their attention in a top-down manner toward different facial expressions (or at least valences), which differ quite subtly in their low-level visual properties. Furthermore, the current study found that when the faces were inverted, participants were no longer reliably able to orient to the cued expressions. Since inversion disrupts the ability to perceive the holistic qualities of faces (Mondloch & Maurer, 2008; Searcy & Bartlett, 1996; Yin, 1969), these findings suggest that participants relied on a holistic top-down set of an emotional face to be able to rapidly shift attention toward it. Surprisingly, however, faster overall responses were found in Experiment 2 in which the faces were inverted, compared with Experiment 1, which employed upright faces. One possibility is that, because participants could not reliably orient to the cued inverted face, they relied on bottom-up capture of attention by the probe instead. This is consistent with the rapid responses that were found in Experiment 3, when no top-down cue was provided. This reveals one potential limitation of the dot-probe task employed in this study; faster RTs may be found when participants were watching out for the probe compared with when orienting to the cued faces. One method of overcoming this issue would

be to have participants answer a question about the target itself (rather than a probe that appeared in the locus of the target). For example, a common design with more basic stimuli has been a singleton search with one target amongst distractors. Typically, the target is defined by its differing shape (e.g., diamond amongst circles) or colour (red circle amongst green circles) and participants are asked to respond to the orientation of a line segment presented inside the target shape (Theeuwes et al., 2006). Therefore, a modification of the current study could be to respond to the gender of the target face.

Selection history appeared to play no role in facilitating attentional orienting to facial expressions in this study. This was supported both by trial-by-trial analyses (Experiment 1) and keeping the cued expression constant across a block of trials (Experiment 4). This contrasts with some previous research which claimed that spatial orienting toward visual stimuli was driven by selection history and not top-down attention (Theeuwes et al., 2006; Theeuwes, 2013). For example, Theeuwes et al. (2006) cued on each trial (with 80% validity) a target, which could either be a shape or colour singleton. No cueing effects were found, indicating that foreknowledge (i.e., top-down attention) did not facilitate attentional orienting. By contrast, when a symbol cue was used, in which the cue was identical to the target, cueing effects were revealed (even when the cue was not predictive of the target). This research demonstrates that top-down attention did not guide attention to featural singletons, as instead, improved performance was driven by selection history (labelled as "bottom-up priming" in this study). Due to concerns that employing a singleton pop-out target made the task too easy, Theeuwes and Van der Burg (2011) conducted a series of similar experiments in which there were two pop-out stimuli (red or green colour singletons) presented alongside five grey distractors. These researchers found that when a verbal cue was used, participants were unable to inhibit attention to an irrelevant coloured distractor, whereas when a symbol cue was used, participants were able to more reliably attend to the target and ignore the distractor.

This is likely because the symbol cue, unlike the verbal cue, primed the features of the target, resulting in cue-target priming (which can be thought of as a type of selection history). Overall, the results of Theeuwes and Van der Burg (2011) and Theeuwes et al. (2006) indicate that, even for difficult displays when there are two featural singletons competing for attention, selection history and not top-down attention drives attentional orienting to features.

However, Belopolsky and Awh (2016) found top-down guidance of attention for visually complex arrays in which pop-out was not present. That is, these researchers used a similar design to Theeuwes et al. (2006) in which participants were presented with a verbal colour cue and then viewed a six-item display and were asked to respond to the orientation of the line segment contained in the target. When a pop-out design was used, in which the target was a separate colour to the distractors, selection history effects were revealed but top-down knowledge of the target did not reliably improve performance. By contrast, if the distractors each had a different colour in the display, and so pop-out effects were eliminated, both selection history and top-down attention had separate contributions to target detection. Other research has confirmed the role of top-down attention to spatial attentional orienting using basic stimuli (Becker, 2018; Becker, Lewis, & Axtens, 2017).

The absence of a selection history effect in the present study indicates that we may have identified some boundary conditions for selection history. This, in turn, has the potential to provide important information about the nature of selection history. There are three key differences between the present study and the methods that have been used to date to examine selection history. It is possible that one of these factors prevented selection history effects from emerging, or a combination of two or all three of the factors. Firstly, photographs of faces are visually more complex in a number of ways than simple shape and colour cues (e.g., green circle). This raises the possibility that selection history effects are only present for basic stimuli such as simple shapes and colours. If so, then this would be important to know,

because it would inform us that selection history effects are likely to be of little consequence in real-world vision, and therefore should not be a focus of laboratory research. Secondly, in the current study, the same expression either repeated or switched across trials. However, this expression could be from different models (i.e., different individual faces). Thus, selection history effects may be limited to repetition of identical exemplars, and not generalise to category-level effects. This would reflect a dissociation between selection history and top-down attention, which can operate at category level beyond individual exemplars (Goodhew et al., 2014; Most, 2013; Wyble et al, 2013). A straightforward way of testing this explanation and distinguishing it from the first would be to determine whether selection history effects emerge for faces when the same face is presented across trials. Thirdly, selection history may be intrinsically tied to the nature of the task. In standard selection history paradigms participants make a response about the stimulus that is repeated (e.g., indicating the orientation of a line within a red target on two consecutive trials). However, in the present study, the probe to which participants responded was a separate stimulus from the stimuli that they had used to guide their attention (i.e., the faces). It is possible that selection history effects only apply to the actual stimulus that is selected. If so, then they should be observed when participants make a response about the face itself, rather than the stimulus that appears after it. Disentangling these possibilities promises to shed new light on selection history and its function.

Conclusions

The results of the current study are novel in demonstrating the role of top-down attention in orienting toward facial expressions, challenging existing ideas about attentional orienting to facial expressions being driven by bottom-up processes. Specifically, participants could flexibly shift their attention toward both happy and angry faces, even in the presence of competing emotional information. This effect relied on holistic processing rather than

feature-based processing of the faces. Selection history did not affect attentional orienting to faces. This means that we discovered boundary conditions for selection history, highlighting exciting avenues for future research to improve our understanding of the function of selection history. Altogether, the implications of these results are that even when it comes to emotionally-salient facial expressions, individuals are not at the mercy of pre-determined bottom-up salience, but instead, can use their top-down attention in flexible and dynamic ways to orient their attention to goal-relevant stimuli.

Notes

1. This research was supported by an Australian Research Council (ARC) Future

   Fellowship (FT170100021) awarded to S.C.G.

2. Since designing the study, it has come to our attention that G*Power does not allow

   for the computation of power for designs with more than one repeated-measures

   factor, as was present here. When we specify eight measurements G*Power assumes

   that these all reflect levels of one variable (8x1), rather than a 2x2x2 factorial design.

   For future research we plan to use the superior newly available calculators which

   permit these sorts of common experimental designs, in addition to permitting for

   correction for publication bias (Anderson, Kelley, & Maxwell, 2017). However, since

   G*Power was used, its results are reported here.

3. In the raw data deposited in OSF, for **Experiment 1**, the columns are as follows:

   Participant Number = the de-identified participant number each individual was

   assigned in chronological order, *InvalidKeyPress* = the percentage of trials for which

   an invalid key press was made (i.e., a non-response key), which were excluded from

   further analysis, and *Excluded RT* = the percentage of trials excluded due to outlier

   screening (see definitions in main manuscript). *AccOverall* = overall accuracy.

   RTOverall = overall RT. Note that overall accuracy and RT per participant were

   compiled from all correct trials for RT and all trials for overall accuracy (on which a

   valid keypress was made). Note that these differ slightly from the average overall

   accuracy and RT values provided in the results section here, which were calculated by

   taking the mean score of all of the variables. For the next set of columns, any with the

   suffix *Acc* = accuracy data (% correct), and any with the suffix *RT* = response time

   data (in ms). Columns that contain the name *Valid* refer to valid trials whereas *Invalid*

   refer to invalid trials (see main manuscript for definitions). Columns with the name

EmotionNeutral…_Happy = where a happy and a neutral face were presented, and the happy face was cued. *EmotionNeutral…_Angry* = where an angry and a neutral face were presented, and the angry face was cued. *HappyAngry…_Happy* = where a happy and angry face were presented, and the happy face was cued. *HappyAngry_Angry* = where a happy and angry face were presented, and the angry face was cued. *IncorrectCue* = where the cued expression was not shown, and *Happy* = where the probe appeared behind the happy face, and *Angry* = where the probe appeared behind the angry face. For **Experiment 1 Selection History** analysis, columns that begin with an *R* = repeat trials, and *S* = switch trials. The next letter in the name indicates the cued facial expression (i.e., *A* = Angry, *H* = happy), the next two letters refer to the two faces expressed (e.g., *AN* = angry and neutral), and *valid or invalid* refers to whether the trial was valid or not, and *ACC* or *RT* to whether the data are accuracy or RT values. **Experiment 2,** *excludedtotal* = % of trials that met exclusion criteria for invalid responses or RT outlier exclusion, *AccOverall* and *RTOverall* = as per Experiment 1. *HN* = happy and neutral face presented, and happy face cued, *AN* = angry and neutral face presented, angry face cued, *HA* = happy and angry face presented, happy face cued, *AH* = angry and happy face presented, angry face cued. Valid/Invalid and ACC/RT as per Experiment 1. For **Experiment 3,** initial columns as per previous experiments. Then the first two letters refer to the two faces presented, the third letter after the underscore refers to which face the probe appeared behind. For example, HN_N = happy and neutral face presented, probe behind the neutral face. For **Experiment 4,** Condition = whether the participants first completed the Intermixed or Blocked condition. The next set of columns are for the Intermixed condition, and the labels correspond to those of Experiment 2. For example, HAvalidRT = RT for when Happy and Angry face shown, happy face cued, probe

appeared in cued location (valid) in the Intermixed condition. Then the same conditions are shown that arose from the Blocked condition. For instance, Block_HNvalidACC = accuracy percentage for when happy and neutral faces are shown, and the happy face is cued, and the trial is valid from the Blocked Condition.

4. We also considered the role of participant handedness in these analyses (we thank an anonymous reviewer for this suggestion). All of the effects reported here in Experiment 1 were unchanged (i.e., stayed significant / non-significant) when the analyses focussed solely on right-handed participants, with the exception of the trend toward an interaction between target face and distractor face, which became significant ($p = .041$). However, the key result – the main effect of Cue Validity – was unchanged. Similarly, in Experiment 2, the results of the ANOVA on RT were unchanged by excluding non-right-handed participants (all remained non-significant, ps > =.131). Furthermore, the absence of an effect of Condition on RT in Experiment 3 remained even for the exclusively right-handed sample ($p = .321$). Finally, in Experiment 4, excluding non-right-handers made the effect of Cue Validity trend-level ($p = .087$), while the main effect of Condition remained significant ($p = .001$). The crucial interaction between Condition and Cue Validity remained non-significant ($p = .114$). All other effects remained non-significant.

5. Surprisingly though, a significant four-way interaction was found between target expression, paired face, validity, and repetition, $F(1, 42) = 4.37$, $p = .043$, $\eta^2 = .094$. To further analyse this effect, an ANOVA was conducted for the happy and angry target faces separately. This found that, for the angry faces, the interaction between paired face, validity, and repetition were not significant, $F(1, 42) = 2.04$, $p = .161$, $\eta^2 = .046$. This indicates that, when participants were orienting to angry faces (when paired with either a neutral or happy face), there was no effect of selection history.

For the happy target face, the interaction between the paired face, validity and repetition was also not significant, $F(1, 42) = 3.25$, $p = .079$, $\eta^2 = .072$. These results indicate that no clear evidence was found for selection history.

6. Note that additional preliminary analyses were conducted for sequence effects (whether participants completed the mixed or blocked condition first). An interaction was found between condition and sequence effects (with large reductions in RT found in the mixed condition when it was performed second in sequence, compared with when it was performed first). However, these effects did not interact with validity (*ps* ≥ .229).

References

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more

    accurate statistical power: A method adjusting sample effect sizes for publication bias

    and uncertainty. *Psychological Science, 28*(11), 1547-1562.

    doi:10.1177/0956797617723724.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional

    control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437-443.

    doi:10.1016/j.tics.2012.06.010

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M.

    H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A

    meta-analytic study. *Psychological Bulletin, 133*(1), 1-24. doi:10.1037/0033-

    2909.133.1.1

Barratt, D., & Bundesen, C. (2012). Attentional capture by emotional faces is contingent on

    attentional control settings. *Cognition and Emotion, 26*(7), 1223-1237.

    doi:10.1080/02699931.2011.645279

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The

    face in the crowd effect unconfounded: Happy faces, not angry faces, are more

    efficiently detected in single- and multiple-target visual search tasks. *Journal of

    Experimental Psychology: General, 140*(4), 637-659. doi:10.1037/a0024060

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The

    confounded nature of angry men and happy women. *Journal of Personality and

    Social Psychology, 92*(2), 179-190. doi:10.1037/0022-3514.92.2.179

Becker, S. I. (2018). Reply to Theeuwes: Fast feature-based top-down effects, but saliency

    may be slow. *Journal of Cognition, 1*(1), 28, 1-3. doi:10.5334/joc.23

Becker, S. I., Lewis, A. J., & Axtens, J. E. (2017). Top-down knowledge modulates onset capture in a feedforward manner. *Psychonomic Bulletin and Review, 24*(2), 436-446. doi:10.3758/s13423-016-1134-2

Belopolsky, A. V., & Awh, E. (2016). The role of context in volitional control of feature-based attention. *Journal of Experimental Psychology: Human Perception and Performance, 42*(2), 213-224. doi:10.1037/xhp0000135

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436. doi:10.1163/156856897X00357

Broadbent, D. E. (1958). *Perception and communication*. New York, NY: Pergamon Press.

Büsel, C., Voracek, M., & Ansorge, U. (2018). A meta-analysis of contingent-capture effects. *Psychological Research*. doi:10.1007/s00426-018-1087-3

Corbetta, M. & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-215. doi:10.1038/nrn755

Cox, J. A., Christensen, B. K., & Goodhew, S. C. (2017). Temporal dynamics of anxiety-related attentional bias: is affective context a missing piece of the puzzle? *Cognition & Emotion, 32*(6), 1329-1338 doi:10.1080/02699931.2017.1386619

Delchau, H. L., Christensen, B. K., O'Kearney, R., & Goodhew, S. C. (in press). What is top-down about seeing enemies? Social anxiety and attention to threat. *Attention, Perception, & Psychophysics*. doi:10.3758/s13414-019-01920-3

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222. doi:10.1146/annurev.ne.18.030195.001205

Folk, C. L., & Remington, R. (2010). A critical evaluation of the disengagement hypothesis. *Acta Psychologica, 135*(2), 103-105. doi:10.1016/j.actpsy.2010.04.012

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology:*

Human Perception and Performance, 18(4), 1030-1044. doi:10.1037//0096-1523.18.4.1030

Goodhew, S. C., Kendall, W., Ferber, S., & Pratt, J. (2014). Setting semantics: Conceptual set can determine the physical properties that capture attention. *Attention, Perception and Psychophysics, 76*(6), 1577-1589. doi:10.3758/s13414-014-0686-3

Grafton, B., & MacLeod, C. (2016). Engaging with the wrong people: The basis of selective attention to negative faces in social anxiety. *Clinical Psychological Science*, *4*(5), 793-804. doi:10.1177/2167702615616344

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54*(6), 917-924. doi:10.1037/0022-3514.54.6.917

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior, 21*(4), 241-257. doi:10.1023/A:1024952730333

Horstmann, G., Lipp, O. V., & Becker, S. I. (2012). Of toothy grins and angry snarls - Open mouth displays contribute to efficiency gains in search for emotional faces. *Journal of Vision, 12*(5): 7, 1-15. doi:10.1167/12.5.7

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes Factors. *Journal of Problem Solving, 7*(1), 2-9. doi:10.7771/1932-6246.1167

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219-227. doi:10.1007/978-94-009-3833-5_5

Lin, M., Hofmann, S. G., Qian, M., Kind, S., & Yu, H. (2016). Attention allocation in social

   anxiety during a speech. *Cognition and Emotion, 30*(6), 1122-1136.

   doi:10.1080/02699931.2015.1050359

Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009a). Emotional faces in neutral crowds:

   Detecting displays of anger, happiness, and sadness on schematic and photographic

   images of faces. *Motivation and Emotion, 33*(3), 249-260. doi:10.1007/s11031-009-

   9136-2

Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009b). No effect of inversion on attentional and

   affective processing of facial expressions. *Emotion, 9*(2), 248-259.

   doi:10.1037/a0014715

LoBue, V., Rakison, D. H., & DeLoache, J. S. (2010). Threat perception across the life span:

   Evidence for multiple converging pathways. *Current Directions in Psychological

   Science, 19*(6), 375-379. doi:10.1177/0963721410388801

Macleod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders.

   Journal of Abnormal Psychology, 95(1), 15-20. doi: 10.1037//0021-843x.95.1.15

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory and

   Cognition, 22*(6), 657-672. doi:10.3758/BF03209251

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour

   Research and Therapy, 36*(9), 809-848. doi:10.1016/S0005-7967(98)00063-1

Mondloch, C. J., & Maurer, D. (2008). The effect of face orientation on holistic processing.

   Perception, 37(8), 1175-1186. doi:10.1068/p6048

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau

   (2005). *Tutorials in Quantitative Methods for Psychology, 4*(2), 61-64.

   doi:10.20982/tqmp.04.2.p061

Most, S. B. (2013). Setting sights higher: Category-level attentional set modulates sustained

    inattentional blindness. *Psychological Research, 77*(2), 139-146. doi:10.1007/s00426-

    011-0379-7

Neisser, U. (1967). *Cognitive Psychology.* East Norwalk, CT, US: Appleton Century Crofts.

Öhman, A. (2007). Has evolution primed humans to "beware the beast"? *Proceedings of the*

    National Academy of Sciences of the United States of America, 104(42), 16396-

    16397. doi:10.1073/pnas.0707885104

Öhman, A & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved

    module of fear and fear learning. *Psychological Review, 108*(3), 483-522.

    doi:10.1037/0033-295X.108.3.483

Pishyar, R., Harris, L. M., & Menzies, R. G. (2004). Attentional bias for words and faces in

    social anxiety. *Anxiety, Stress and Coping, 17*(1), 23-36.

    doi:10.1080/10615800310001601458

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental*

    Psychology, 32(1), 3-25. doi:10.1080/00335558008248231

Purcell, D. G., Stewart, A. L., & Skov, R. B. (1996). It takes a confounded face to pop out of

    a crowd. *Perception, 25*(9), 1091-1108. doi:10.1068/p251091

Roelofs, K., Hagenaars, M. A., & Stins, J. (2010). Facing freeze: Social threat induces bodily

    freeze in humans. *Psychological Science, 21*(11), 1575-1581.

    doi:10.1177/0956797610384746

Savage, R. A., Lipp, O. V., Craig, B. M., Becker, S. I., & Horstmann, G. (2013). In search of

    the emotional face: Anger versus happiness superiority in visual search. *Emotion,*

    13(4), 758-768. doi:10.1037/a0031970

Schoeberl, T., Goller, F., & Ansorge, U. (2019). Testing a priming account of the contingent-

    capture effect. Attention, Perception, & Psychophysics. doi:10.3758/s13414-019-

    01672-0

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-

    relational information in faces. *Journal of Experimental Psychology: Human*

    Perception and Performance, 22(4), 904-915. doi:10.1037/0096-1523.22.4.904

TeamJASP. (2018). JASP (Version 0.9).

Theeuwes, J. (2013). Feature-based attention: It is all bottom-up priming. *Philosophical*

    Transactions of the Royal Society B-Biological Sciences, 368(1628): 20130055, 1-11.

    doi:10.1098/Rstb.2013.0055

Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and

    Volitional; A Reply to Commentaries. Journal of Cognition, 1(1), 21.

    doi:10.5334/joc.32

Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No

    top-down modulation, only bottom-up priming. *Visual Cognition, 14*(4-8), 466-489.

    doi:10.1080/13506280500195110

Theeuwes, J., & Van der Burg, E. (2007). The role of spatial and nonspatial information in

    visual selection. *Journal of Expimental Psychology: Human Perception and*

    Performance, 33(6), 1335-1351. doi:10.1037/0096-1523.33.6.1335

Theeuwes, J., & Van der Burg, E. (2011). On the limits of top-down control of visual

    selection. *Attention, Perception, & Psychophysics, 73*(7), 2092-2103.

    doi:10.3758/s13414-011-0176-9

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A.,...Nelson, C.

    (2009). The NimStim set of facial expressions: Judgments from untrained research

participants. *Psychiatry Research*, *168*(3), 242–249.

doi:10.1016/j.psychres.2008.05.006

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248. doi:10.1080/17470216008416732

Wells, L. J., Gillespie, S. M., & Rotshtein, P. (2016). Identification of emotional facial expressions: Effects of expression, intensity, and sex on eye gaze. *PLoS One, 11*(12). doi:10.1371/journal.pone.0168307

Williams, M. A., Moss, S. A., Bradshaw, J. L., & Mattingley, J. B. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual Cognition, 12*(1), 29-50. doi:10.1080/13506280444000193

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1(2), 202-238. doi:10.3758/BF03200774

Wyble, B., Folk, C., & Potter, M. C. (2013). Contingent attentional capture by conceptually relevant images. *Journal of Experimental Psychology: Human, Perception and Performance, 39*(3), 861-871. doi:10.1037/a0030517

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141-145. doi:10.1037/h0027474

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the Cross-Race Effect. *Personality and Social Psychology Review, 16*(2), 116-142. doi:10.1177/1088868311418987